Master of Public Health

(European Public Health Master)

Master de Santé Publique

Income inequality and disease burden in OECD countries: a hierarchical clustering and panel data analysis with Two-Way Fixed Effects

Berta PIQUÉ SMITH

Europubhealth 2023-2025, Master II

Specialisation: Epidemiology & Biostatistics

Location of the practicum: Organization for the Economic Cooperation and Development (OECD), Paris, France

Professional advisor: Eliana Barrenho

Academic advisor: Tarik Benmarhnia

Acknowledgements

Writing this master's thesis have been much better with all the support network that have helped me throughout the entire process. It is to these people that I want to sincerely express my gratitude for their unconditional help.

First of all, I would like to thank Eliana Barrenho, Policy Analyst at the Pharmaceutical and Medical Devices team at the OECD, and the entire team from which I discovered the field of Health Economics. Thank you for your guidance and for sharing invaluable insights from which I have learned so much. I also want to thank Isaura, because your encouragement during our calls boosted my morale and made me believe we could overcome any biostatistical challenge.

To Professor Jay Kaufman, because despite the distance, you were always available to discuss and guide me whenever I needed it. Your insights came at a moment when I was lost and helped me unblock my work. I am deeply grateful.

To Mélanie Bertain, because your support has been essential in the writing of this thesis. Thank you for your patience and for helping me refine my ideas.

To my family, who have listened even when the topic was far from your own experiences. Thank you for everything you have taught me; without you, I would not be who I am.

To my friends, because you are an essential part of my live and you have the ability to make any problem feel like no problem at all.

Finally, to Tere, without question, you have been the person who has heard me talk about this thesis more than anyone else. Here's to our team, which can achieve anything!

Table of contents

List of acronyms **3**

Abstract **4**

Résumée **5**

1. Introduction **6**

1.1. Historical context of health inequities 6

1.2. Measuring income inequality: The Gini Coefficient and Lorenz Curve 7

1.3. Literature gaps 9

1.4. Justification of the study and objectives 9

2. Methodology **10**

2.1. Sample design 10

2.2. Data sources and data collection procedures 10

2.3. Variables 11

2.4. Data preparation and imputation 13

2.5. Clustering analysis 13

2.6. Fixed effects panel regression model 15

2.7. Robustness checks and sensitivity analysis 16

2.8. Ethical considerations 16

3. Results **16**

3.1. Clustering analysis 16

3.2. The Gini index and the burden of disease 18

4. Discussion **24**

5. Conclusion **30**

6. References **31**

7. Appendix. Supplementary materials **37**

7.1. Supplementary Material 1 37

7.2. Supplementary Table 1. Descriptive information on variables, data sources and missingness percentage used across 38 countries from 2009 to 2019 38

7.3. Supplementary Figure 1. Importance of variables in defining the clusters measured by the Mean Decrease Gini (a) and the Mean Decrease accuracy (b) 41

7.4. Supplementary Table 2. Top five diseases driving cluster separation in the random-forest analysis 42

7.5. Supplementary Table 3. Age-standardised DALY rates for the 24 leading diseases across the three epidemiological clusters identified among OECD countries 43

7.6. Supplementary Table 4. List of countries according to their cluster 46

7.7. Supplementary Figure 2. World Map of Gini index in 2009 with OECD countries coloured ... 47

7.8. Supplementary Figure 3. World Map of Gini index in 2019 with OECD countries coloured ... 48

7.9. Supplementary figure 4. Evolution of the Gini index from 2009 to 2019 49

7.10. Supplementary Table 5. Descriptive information on socioeconomic indicators and care services in 2009 and 2019 in the 38 OECD countries 50
7.11. Supplementary Table 6. Spearman correlation between the studied covariates 52
7.12. Supplementary Table 7. Country and year fixed effect crude model..... 53
7.13. Supplementary Table 8. Country and year fixed effect model using with Palma ratio 54

List of acronyms

ASDR	Age-standardized DALY rate
DALY	Disability-adjusted life year(s)
GBD	Global Burden of Disease
GDP	Gross domestic product
GHO	Global
IHME	Institute for Health Metrics and Evaluation
IHD	Ischaemic Heart Disease
IQR	Inter-quartile range
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary least squares
RF	Random Forest
SD	Standard deviation
SDI	Socio-demographic Index
SE	Standard error
SDG	Sustainable Development Goal
UN	United Nations
VIF	Variance inflation factor
YLL	Years of life lost
YLD	Years life with disability
WHO	World Health Organization

Abstract

Background: income inequality can have a harmful impact on population health explained by underinvestment in social goods, such as health care and public education, reduction of social cohesion and social capital, and through exposing individuals to detrimental psychosocial effects. However, there are some discrepancies as there is a body of the literature that shows no association between income distribution and health outcomes, and there is limited evidence from multi-country studies. This study investigates the association between income inequality and disability-adjusted life years (DALYs) in the 38 OECD countries from 2009 to 2019 and explores whether behavioural and metabolic risk factors modify this relationship.

Methods: data were aggregated across genders and obtained from publicly available datasets comprising the Global Burden of Disease 2021, the Global Health Observatory, the Organization for the Economic Cooperation and Development, the World Bank and the World Income Inequality Database, with a total of 418 country-year observations. Hierarchical clustering of OECD countries based on their epidemiological profiles was conducted to select one representative disease per cluster. Fixed-effects regression models explored the association between Gini index and DALY rates for those diseases, adjusting for health system and macroeconomic variables. Metabolic and behavioural risk factors were also included.

Findings: a one-point increase in the Gini index was associated with an increase in prenatal preterm birth of 1.5% (SE = 0.002) DALY rates and a decrease in ischaemic heart disease (IHD) of 1.3% (SE = 0.005) DALY rates when adjusting for covariates. These associations weakened slightly after accounting for risk factors.

Conclusion: this ecological study reveals that, across the 38 high-income OECD countries, the association between income inequality and DALY rates is heterogeneous and disease specific. These results highlight the need for disease-specific policies to address health inequalities.

Keywords: income inequality, Gini index, disease, DALY.

Résumée

Contexte: les inégalités de revenus peuvent avoir un effet néfaste sur la santé de la population en raison du sous-investissement dans les biens sociaux, tels que les soins de santé et l'éducation publique, de la réduction de la cohésion sociale et du capital social, et de l'exposition des individus à des effets psychosociaux néfastes. Toutefois, il existe certaines divergences, car une partie de la littérature ne montre aucune association entre la répartition des revenus et les résultats en matière de santé, et les preuves provenant d'études multi-pays sont limitées. Cette étude examine l'association entre les inégalités de revenus et les années de vie corrigées de l'incapacité (DALY) dans les 38 pays de l'OCDE de 2009 à 2019 et cherche à savoir si les facteurs de risque comportementaux et métaboliques modifient cette relation.

Méthodes: les données ont été agrégées par sexe et obtenues à partir d'ensembles de données accessibles au public comprenant la Global Burden of Disease 2021, l'Observatoire mondial de la santé, l'Organisation de coopération et de développement économiques, la Banque mondiale et la base de données sur les inégalités de revenus dans le monde, avec un total de 418 observations par année-pays. Un regroupement hiérarchique des pays de l'OCDE basé sur leurs profils épidémiologiques a été effectué afin de sélectionner une maladie représentative par regroupement. Des modèles de régression à effets fixes ont exploré l'association entre l'indice de Gini et les taux de DALY pour ces maladies, en tenant compte du système de santé et des variables macroéconomiques. Les facteurs de risque métaboliques et comportementaux ont également été pris en compte.

Résultats: une augmentation d'une unité de l'indice de Gini est associée à une augmentation de la prématurité prénatale de 1,5 % (SE = 0,002) du taux de DALYs et à une diminution des cardiopathies ischémiques de 1,3 % (SE = 0,005) du taux de DALYs après ajustement en fonction des covariables liées au système de santé et au financement de la santé. Ces associations se sont légèrement affaiblies après prise en compte des facteurs de risque.

Conclusion: cette étude écologique révèle que, dans les 38 pays de l'OCDE à revenu élevé, l'association entre les inégalités de revenu et les taux de DALY est hétérogène et spécifique à la maladie. Ces résultats soulignent la nécessité de mettre en place des politiques spécifiques à chaque maladie pour lutter contre les inégalités en matière de santé.

Mots-clés: inégalités de revenus, indice de Gini, maladie, DALY.

1. Introduction

1.1. Historical context of health inequities

Health is influenced by different social, economic, and political factors that extend far beyond the scope of biology and other determinants at the individual level. Overall population health has improved over the past two centuries due to industrialization and biomedical innovation (1,2), however, these health benefits have not been equally distributed across and within countries. The study of Ian Gregory (2009) showed that despite the medical, social and economic changes, the relation between poverty and mortality across England and Wales remains as strong today as it was at the start of the 20th century (3), reflecting the persistent influence of structural determinants on health outcomes. Regarding life expectancy, recent data from OECD show substantial education-related disparities in life expectancy. On average, the gap at age 25 between high- and low- educated groups is 8.2 years for men and 5.2 years for women (4).

Determinants of health are complex and intrinsically related to each other, being social class one of the most widely studied. As early as the 17th century, demographers such as William Petty documented associations between social divisions and health (5). Later, in the nineteenth century, the pathologist Rudolf Virchow connected a typhus outbreak with economic depression, famine and poor living conditions of the population (6). More recently, the Whitehall study of British civil servants reported an inverse association between occupational grade and both mortality and morbidity (7).

Beyond socioeconomic factors, Dahlgren and Whitehead (1991) presented the "Social determinants of health" framework, which considers a broader range of health determinants and incorporates additional actors such as social support networks, community resources, and access (8). As Wilkinson and Marmot (2003) argue, lifestyle behaviours are closely linked to the availability of choices shaped by an individual's socioeconomic environment (9). In this line, the study by Dumas et al. (2014) based on interviews with underprivileged young women in low-income areas of Québec, Canada, found that obese women facing socioeconomic difficulties were not engaging with weight management strategies. The main reason was that they prioritized imminent necessities, including economic stability, taking care of family needs or healing from highly debilitating illnesses. Also, some women stated a lack of free time or resources to pay for a gym and noted that healthy foods were often the most expensive (10). This lack of choice could be the reason why behavioural risk factors including alcohol, tobacco, physical inactivity and poor diet have been strongly associated with lower socioeconomic status (11,12).

The COVID-19 pandemic has further increased health inequalities as the most vulnerable populations have been the most affected (13). As inequality remains a major challenge, the United Nations has established the Sustainable Development Goal 10 which aims to “reduce inequality within and among countries”. This highlights that equity is not only a matter of social justice but also an essential prerequisite for achieving sustainable development (14).

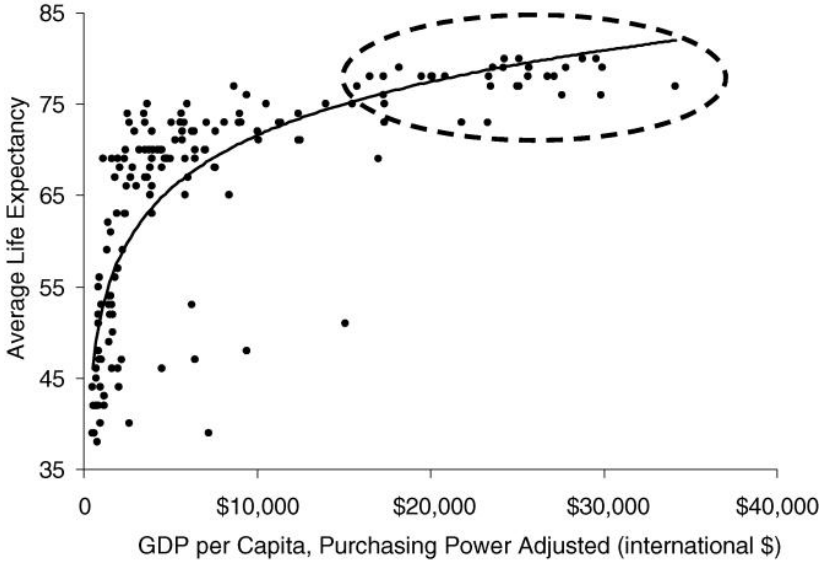
1.2. Measuring income inequality: The Gini Coefficient and Lorenz Curve

Social scientists have agreed that society is much more than the simply sum of individuals, and the factors determining population health cannot be reduced to individual-level risk factors alone (15). In this context, the income inequality theory has emerged as a central concept within the broader field of social determinants of health. It proposes two main pathways through which inequality can lead to health disparities: the neo-materialist and the psychosocial.

The neo-materialistic pathway argues that material conditions, such as inadequate housing, poor nutrition or limited access to clean water have a direct effect on health. However, the sociologists Marmot and Wilkinson state that even when the basic needs are covered, the psychosocial effects of relative deprivation including insecurity, anxiety and social isolation, among others, can still undermine population health (16). These effects are shaped by an individual’s position within the broader socioeconomic structure.

This theoretical distinction is illustrated by the nonlinear relationship between gross domestic product (GDP) and life expectancy (Figure 2). While rising incomes improve life expectancy in low-income countries, this effect plateaus in high-income economies, implying that beyond a certain threshold, factors such as income distribution and other social determinants may become more influential for population health rather than the absolute national wealth (15).

Figure 2. GDP per capita distribution and average life expectancy



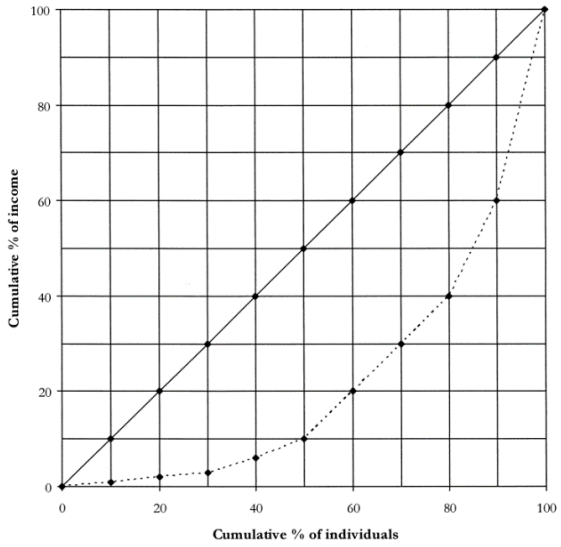
Source: World Bank 2002.

Growing evidence supports this psychosocial interpretation, as studies have shown that societies with higher income inequality often exhibit lower levels of community engagement, social cohesion and trust, elements collectively referred to as social capital (22). As described by Putnam, social capital refers to features of social organizations, such as networks, norms and trust that facilitate action and cooperation for mutual benefit (17).

Taken together, these insights have prompted researchers to look beyond average national income and examine the effect of income distribution as a more meaningful determinant health. The income inequality hypothesis thus provides a framework for understanding how both material and psychosocial mechanisms contribute to unequal health outcomes across and within countries (18).

While several inequality measures exist, including the Theil index, Palma ratio, and Atkinson index, the Gini coefficient remains the most widely used indicator of income inequality (19). Mathematically, the Gini index is defined as half of the arithmetic average of the absolute differences between all pairs of incomes in a population, the total then being normalized on mean income. It ranges from 0 representing perfect equality, to 1 under perfect inequality and is illustrated using a *Lorenz curve* (Figure 1). On the horizontal axis, households are ranked from the lowest decile income group to the top decile group and the vertical axis represent the proportion of the aggregate income each group earns. Under perfect equality, the Lorenz curve aligns with a 45-degree line of equality; deviations from this line represent growing inequality. For example, as shown in Fig. 2, if the bottom 50% of households earn only 10% of total income, the curve falls significantly below the equality line (20).

Figure 1. Lorenz curve



Source: Subramanian SV and Kawachi I. et al 2004

1.3. Literature gaps

In recent decades, epidemiological studies have employed different statistical models to investigate the effects of income inequality on population health. However, whether income inequality is a hazard to public health is still a matter of debate, with different studies showing no significant association between income distribution and health outcomes (15,21–23).

Even though the research on income inequality and health is vast, the focus has predominantly been on life expectancy and mortality, with fewer studies examining broader metrics such as disability-adjusted life years (DALYs). However, DALYs capture both premature mortality and years lived with disability, offering a broader measure of burden of disease. When DALYs are used, they are typically aggregated for broad disease categories, rather than being considered by specific health conditions (24,25). Furthermore, most of the existing literature relies on subnational data, focusing on disparities within specific regions, cities, or states from single countries such as the United States (US) (26), Korea (27) and Italy (28) among others, with relatively few using longitudinal, within-country designs across multiple high-income settings.

The discrepancies found in the literature combined with the complexity of studying income inequalities, as they are highly entangled with other social determinants of health, highlight the need for a broader scope.

1.4. Justification of the study and objectives

The aim of this study was to assess the association between income inequality, measured by the Gini index, and disease burden, measured by age-standardized DALY rates (ASDR) per 100,000, at the disease-country-year level in countries of the OECD from 2009 to 2019.

Considering the income inequality theory, which postulates that population health is influenced by income distribution, we hypothesized that increases in the Gini index would be associated with higher ASDR for the studied diseases. Our secondary hypothesis postulated that behavioural and metabolic risk factors may modify the relationship observed between inequality and health.

The objectives of this thesis are:

1. To conduct a clustering analysis of the 38 OECD countries based on their ASDR for the 10 leading causes of DALYs in each country, in order to map the different epidemiological profiles across OECD nations.
2. To explore the association between inequality and ASDR for the diseases previously selected, controlling for health system financing variables such as healthcare

expenditure and public health coverage, for health workforce and access indicators proxied by physician and bed density and for unemployment rate.

3. To assess whether behavioural and metabolic risk factors are associated with or modify the relationship between income inequality and ASDR for the selected diseases.

The empirical analysis relies on a longitudinal panel dataset from 2009 to 2019 and uses two-way fixed-effects models to estimate within-country variation over time, accounting for unobserved heterogeneity. By combining longitudinal analysis with clustering driven disease selection, the study seeks to provide a more contextualized understanding of how structural determinants, focusing on income inequality, may be related to variation in disease burden in different national settings.

This master's thesis was conducted during my internship at the Organization for Economic Co-operation and Development (OECD), within the Health Division and specifically in the Pharmaceuticals and medical devices team. During the internship, I contributed to the drafting of a policy report on pharmaceutical research and development and unmet medical needs. I carried out descriptive and quantitative analysis using different statistical methodologies with the Global Burden of Disease data. While this report was separate from the thesis project, the insights gained were important in shaping the analytical approach of this work. The thesis was developed independently, under the supervision of my OECD professional advisor.

2. Methodology

2.1. Sample design

This longitudinal ecological study followed the 38 members countries of the OECD[1] over the period from 2009 to 2019. The dataset is a balanced panel of 418 country-year observations. To explore associations within countries over time while accounting for the hierarchical nature (years nested within countries) of the data panel regression model was used. Panel data models with two-way fixed effects, at the year and country level, were employed to examine the within-country effects of income inequality over time while accounting for unobserved heterogeneity across countries and years.

2.2. Data sources and data collection procedures

The information regarding ASDR for the selected diseases was extracted from the Global Burden of Disease (GBD) 2021 data (29), generated by the Institute for Health Metrics and

[1] The OECD countries include Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, South Korea, Latvia, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, Türkiye, the United Kingdom, and the United States.

Evaluation (IHME). The GBD was chosen as it is the most comprehensive epidemiological global source for prevalence and mortality data that allows country comparison. GBD 2021 uses the latest epidemiological data and standardized methods to assess health outcomes including incidence, prevalence, mortality, years of life lost (YLL), years lived with disability (YLD) and DALYs, with different measures across 369 diseases, injuries, and impairments and 87 risk factors by age and sex for 204 countries and regions. The methodology applied in the GBD 2021 has been described in detail in other materials (29,30). Data for ASDR were collected from IHME's publicly available VizHub tool for the years 2009 to 2019 (31) and data for the covariates included in the regression model were collected from publicly available datasets comprising the Global Health Observatory by the World Health Organization (WHO)(32), the Data Explorer by the OECD (33), the Open Data by the World Bank (34) and the World Income Inequality Database from the United Nations (35). Supplementary Table 1 presents descriptive information about the variables and data sources used together with the percentage of missingness.

We acknowledge that although the IHME, in collaboration with the WHO, provides regularly updated estimates, these are modelled rather than directly observed data and may be different from national statistics. In addition, GBD do not capture the burden of rare and ultra-rare diseases or small population subgroups (36).

2.3. Variables

2.3.1. Dependent Variables

Disease burden was measured with ASDRs per 100,000 population for both sexes aggregated across OECD countries from the GBD database by IHME. The GBD classifies diseases using a four-level hierarchy. Level 1 distinguishes broad categories: non-communicable diseases and injuries, and a combined group of communicable diseases including infectious diseases, maternal and neonatal disorders, and nutritional deficiencies. Level 2 breaks these down into 22 aggregated categories, such as cardiovascular diseases or respiratory infections. Level 3 specifies individual causes like tuberculosis, stroke, or road injuries. For most Level 3 diseases, a more granular Level 4 classification is available. In this study, the most detailed available classification of disease was used. Where Level 4 data existed, Level 3 conditions were disaggregated accordingly, allowing for a more specific description of the disease (30). With the aim of generating epidemiological clusters within OECD countries, the conditions primarily caused by injuries, such as war conflict or firearm injury, were excluded. As a result, we considered 36 level 3 diseases that were disaggregated to 150 level 4 diseases. In summary, data for 150 Level 4 diseases and 158 Level 3 diseases were collected for the 38 OECD countries for the period 2009-2019.

One DALY represents the loss of the equivalent of one year of full health. DALYs for a disease are the sum of the years of life lost to due to premature mortality (YLLs) and the years lived with disability (YLDs). The GBD calculates the YLLs by multiplying cause-age-sex-location-year-specific deaths by the standard life expectancy at the age that death occurred, and the YLDs by multiplying cause-age-sex-location-year-specific prevalence of sequelae by their respective disability weights, for each disease (30). The calculations for DALYs are provided in the Supplementary Material 1. In this study, DALYs were used because they provide a comprehensive and comparable metric of total burden of disease both for fatal and non-fatal health conditions across countries.

For the clustering analysis, the top 10 diseases in ASDR for each OECD country were included which led to the analysis of a total of 24 diseases, as leading causes were not the same across countries. For the subsequent regression analysis, one disease per cluster was chosen based on the following rationale: 1) present a statistically significant Kruskal–Wallis test of median ASDR differences across clusters ($p < 0.05$), 2) exemplify the cluster's dominant epidemiological profile and 3) exhibit the largest relative differences in ASDR between clusters. Based on this rationale, the selected diseases were anxiety disorders to represent cluster 1, IHD selected for cluster 2 and neonatal preterm birth for cluster 3. Variable importance was assessed using the Mean Decrease Gini and the Mean Decrease Accuracy and served as a post-hoc check to understand the drivers of clustering, but the results were not used to inform disease selection.

2.3.2. Main explanatory variable

The main independent variable was the country-level household income Gini index, also known as the Gini coefficient. It is a measure designed to quantify the level of inequality within a distribution, in this study it was used to assess income inequality within national populations (37). It ranges from 0 to 100, where 0 indicates perfect income equality and 100 perfect inequality (38). Conceptually, the highest Gini value reflects a scenario in which one household receives all income while everyone else receives none. This variable was used to capture temporal changes in country income distribution.

2.3.3. Control variables

To explore the association between income inequality and the outcome, we adjusted for a set of covariates derived from the OECD Rethinking Health System Performance Assessment framework (39), which maps indicators to the major performance dimensions of health systems, an approach that aligns with the covariates most frequently used in cross-national health inequality research. The covariates included were healthcare expenditure per capita as percentage of GDP (40); the number of hospital beds and the number of physicians per 1,000

inhabitants to account for access to healthcare; the share of health expenditure covered by government and compulsory schemes as percentage of expenditure on health, as a proxy for universal public health coverage, and the national unemployment rate (41). Supplementary Table 1 presents descriptive information about the variables and data sources used.

2.3.4. Risk factors

Risk factors included in the regression models consisted of two behavioural risk factors including the age-standardized prevalence of daily smoking in individuals aged 15 years old and older and the age-standardized prevalence of insufficient physical activity among adults aged 18 years and older. Two metabolic variables, specifically the age-standardized obesity prevalence among adults and the age-standardized prevalence of hypertension among adults aged 30 to 79 years were also included (Supplementary Table 1). These risk factors were selected based on the GBD 2021 comparative-risk assessment which ranked them among the leading modifiable determinants of cardiovascular disease, cancer and overall DALYs burden (42).

2.4. Data preparation and imputation

The dataset was organized as a wide format country-year panel. Missing data was found in 5 of the eleven included variables but the missingness proportion was not over 10% in any of those variables. Patterns of missingness were checked and there was no systematic missingness clustering by country or year, supporting the assumption that values were missing at random. Missing values were single-imputed following OECD rationale for imputation using the nearest observation in time procedure in which values were forward-filled when the next non-missing observation laid ≤ 2 years ahead; otherwise, the nearest prior observation was used.

GDP per capita, poverty (43) and education attainment (44), which are measures widely used in the literature to study health inequalities were excluded due to high collinearity with the main predictor and the other covariates. For the same reason, the risk factors cholesterol, ambient air pollution and alcohol consumption were also omitted.

2.5. Clustering analysis

Ward's hierarchical clustering analysis was performed to group OECD countries into epidemiological profiles using the ASDR for the top 10 diseases per each country, representing a total of 24 diseases due to country differences in the burden of disease. Although all variables were expressed as rates per 100,000 population, their scale differed considerably between diseases, subsequently z-score normalization was used before clustering. Ward's method employs a bottom-up approach where initially each country represents its own cluster, and clusters are iteratively merged until a single cluster is formed. This method ensures the minimum

increase in the total within-cluster sum of squared deviations when two clusters are merged (45). Given that the dataset contained only continuous variables, Euclidean distance was used as the distance metric:

$$Dj = \sum_{j=1}^{K} \left(a_j - b_j\right)^2$$

where *a* and *b* refer to the two observations being compared for the *jth* variable, with *K* as the total number of variables. At each step, the two cases or clusters with the smallest squared Euclidean distance are merged (46).

The dendrogram generated from the hierarchical clustering method provided a visual representation of the clustering structure and the merging process. The Silhouette score, a metric that quantifies how similar an object is to its assigned cluster compared to other clusters was used to define the optimal number of clusters. Silhouette scores range between -1 and 1, with higher values reflecting better-defined clusters. In addition, the optimal number of clusters was confirmed with a visual inspection of the cut corresponding with the partition that demonstrated the greatest relative reduction in inertia.

Additionally, a Random Forest (RF) algorithm was used to evaluate the importance of each variable to the overall clustering structure. This non-parametric learning method uses different decision trees using random subsets of the data to aggregate their outcomes to improve predictive performance. In this study, RF was applied to assess the relative importance of each variable in shaping the resulting cluster structure (Supplementary Figure 1). Variable importance was quantified using two metrics: Mean Decrease Accuracy and Mean Decrease Gini[2] (Supplementary Table 2). The former captures the reduction in model accuracy when a given variable is randomly permuted while keeping all other variables constant, whereas the latter reflects the decrease in node impurity when a variable is used for data splitting within the decision trees. In both cases, higher values reflect a greater contribution to accurate cluster identification. In addition, these two measures were used to assess the specific contribution of each variable to the identification of individual clusters (47).

For each disease, both the mean (± standard deviation, SD) and the median (interquartile range, IQR) were reported to describe the distribution of ASDR within each cluster (Supplementary Table 3). Given the non-normal distribution of the disease burden, the Kruskal–Wallis test was used to assess whether differences across the ASDR cluster medians were statistically significant. To visualize the distribution of disease burden across the identified clusters, boxplots were generated for each of the 24 diseases using raw ASDR per 100,000 population (Figure 1).

[2] Mean Decrease Gini, is a variable-importance metric reported by RF models, is entirely unrelated to the Gini index used to quantify income inequality.

2.6. Fixed effects panel regression model

The data used in this study were panel data of 38 countries from 2009 to 2019. Because the degree of economic development and healthcare system capacity had great heterogeneity across countries, a pooled OLS regression model was excluded in favour of a fixed effect model. Moreover, Hausman test ($p < 0.001$ in all cases) favoured fixed over random effects models (Table 2). Accordingly, country fixed effects were used to control for unobserved factors such as cultural, geographic and historical variables, that vary across countries but are fixed over time, whereas year dummies captured factors that vary over time but are common to all countries. Thus, a two-way fixed effects model was used to simultaneously solve the problem of omitted time invariant variables, but change with country, and vice versa. Two-way fixed effects were estimated using Ordinary Least Square (OLS).

The models were fitted using the ASDR per 100,000 inhabitants as the dependent variable for anxiety disorders, IHD and neonatal preterm birth. Outcomes were log-transformed to satisfy the normality distribution assumption of the residuals. The best transformation was checked using the Boxcox function in R. Whether the model fulfilled the assumptions of no serial autocorrelation and homoskedasticity was checked with the Breusch-Godfrey and Breusch-Pagan tests, respectively. Because residuals exhibited both serial correlation and heteroskedasticity, standard errors were clustered both in time and country. Multicollinearity was evaluated by variance inflation factors (VIF), all of which were below 2.6 in all the models, indicating negligible collinearity.

Six multivariate linear regression models were run, two for each disease. The first one was the adjusted model using the logarithm of ASDR per 100,000 for each disease in a country *i* in year *t* as the dependent variable. The Gini index of country *i* in year *t* as the main predictor and the confounding variables. The α represented country fixed effects, the δ were the year fixed effects and ε was the error term. This is the adjusted model:

$$ln(ASDR_{it}) = \alpha + \beta_1 Gini_{it} + \beta_2 HealthExp_{it} + \beta_3 C_{it} + \alpha_i + \delta_t + \varepsilon_{it}$$

where the C_{it} represented a vector of cofounding variables in country *i* in year *t* including the healthcare expenditure per capita as percentage of GDP, the percentage of unemployment rate among the population aged 25-64, the public health coverage proxied by government and compulsory schemes as a share of total health expenditure and the number of hospital beds and physicians per 1,000 inhabitants.

Subsequently, a second model for each disease explored the adjusted relationship between the Gini index and the ASDR by conditioning on 4 principal risk factors that have been acknowledged to be drivers of global mortality and morbidity (42) and also mediators of the

association between income inequalities and health outcomes[3] (11,48,49). This is the adjusted model including the risk factors:

$$ln(ASDR_{it}) = \alpha + \beta_1 Gini_{it} + \beta_2 HealthExp_{it} + \beta_3 C_{it} + \beta_4 R_{it} + \alpha_i + \delta_t + \varepsilon_{it}$$

where the R_{it} term includes 2 behavioural risk factors, which were daily age-standardized smoking prevalence in individuals aged 15 years old and over and the age-standardized prevalence of insufficient physical activity in adults aged 18 years and older; and 2 metabolic risk factors, consisting of the age-standardized prevalence of obesity among adults aged 18 and older, and age-standardized prevalence of hypertension in adults aged 30 to 79 years.

2.7. Robustness checks and sensitivity analysis

2.7.1. Crude fixed-effect model

To test whether the observed association between Gini index and IHD ASDR remained consistent after removing macro level covariates, each model was re-estimated with only country and year fixed effects and the Gini index (Supplementary Table 7).

2.7.2. Palma ratio as the main predictor variable

A robustness check was carried out in which the Gini index was replaced with the Palma ratio, collected from the World Income Inequality Database from the United Nations (50). This measure is defined as the ratio of the income share of the richest 10% of households to that of the poorest 40%. The results considering both the direction and the statistical significance of the association between ASDR and income inequality remained unchanged (Supplementary Table 8). This confirms that the findings were not driven by the specific inequality metric chosen.

2.8. Ethical considerations

This study is based exclusively on secondary, country-level data obtained from publicly accessible sources; therefore, ethical approval was not required. Ethical issues including plagiarism, misconduct, data fabrication or falsification have been considered by the author.

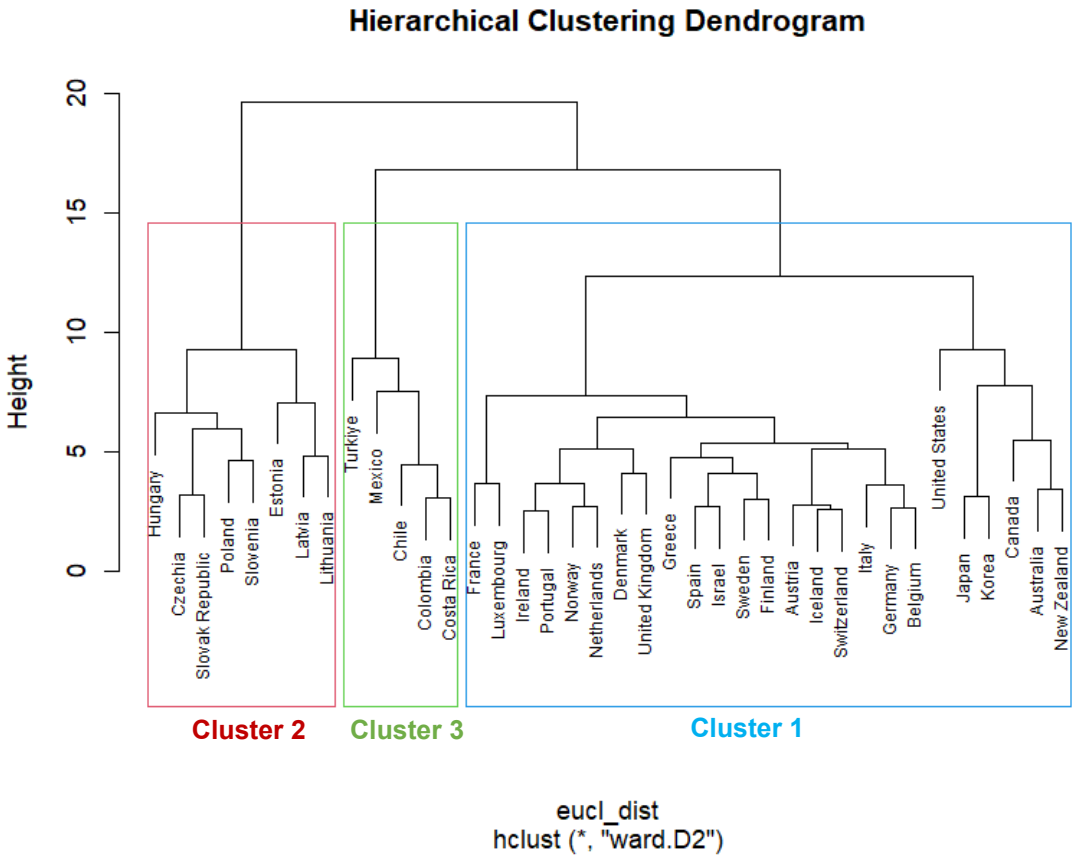
3. Results

3.1. Clustering analysis

Hierarchical clustering of ASDR for the top 10 diseases in individual OECD countries for 2019, covering a total of 24 diseases, partitioned the 38 OECD countries into three epidemiological clusters, as illustrated in the dendrogram (Figure 1).

[3] This refers to previous studies that examined mediation, whereas causality cannot be inferred in the present ecological analysis.

Figure 1. Hierarchical clustering dendrogram for OECD countries with ASDR for 24 diseases

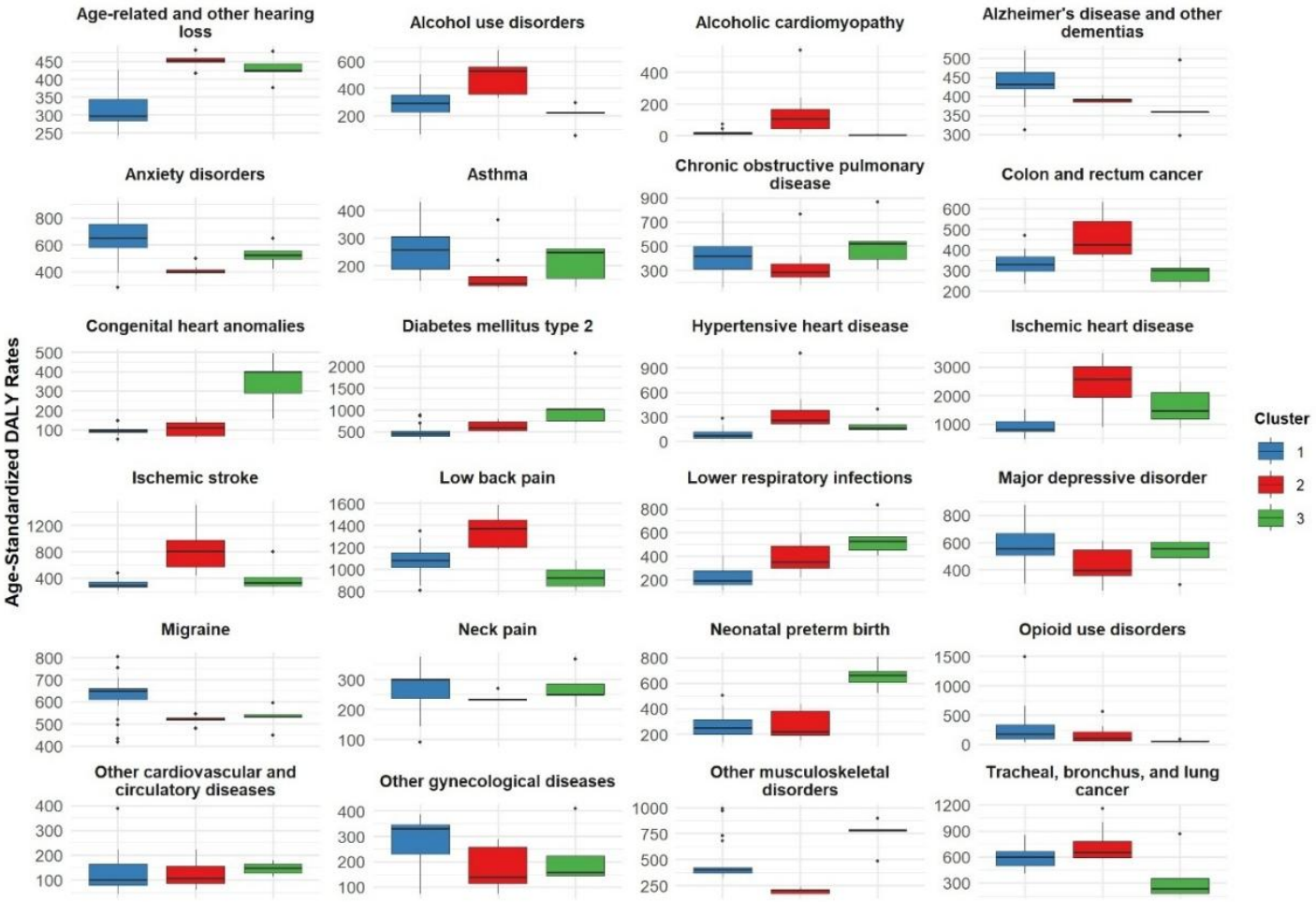


Optimal cluster division measured by the highest silhouette score (silhouette score = 0.29), produced three different epidemiological profiles, including a mental-health and neurodegenerative cluster (n = 25), a cardio-alcoholic cluster (n = 8) and a metabolic-perinatal cluster (n = 5, country list in Supplementary Table 4). Supplementary Table 3 summarises the ASDR medians with the IQR and means with SD for the 24 conditions in each cluster. Inter-cluster differences in medians were tested with the Kruskal-Wallis statistic showing statistically significant differences ($p < 0.05$) for every condition with the exception of chronic obstructive pulmonary disease, major depressive disorder, neck pain, and other cardiovascular and circulatory diseases.

The ASDR for each disease across clusters is shown in Figure 2. This analysis revealed that epidemiological profiles in OECD countries followed geographical distributions. Cluster 1 (n = 25), the mental-health and neurodegenerative cluster, included high-income and developed economies. These countries had the lowest cardiovascular and metabolic burden with the medians of IHD and diabetes mellitus type 2 (T2DM) accounting for ASDR of 802 and 443, respectively. Conversely, this cluster showed the highest burden for mental-health disorders, including anxiety disorders, migraine, and neurodegenerative disorders such as Alzheimer's disease. Cluster 2 (n = 8), grouping Eastern Europe countries, recorded the greatest

cardiovascular burden with the highest ASDR medians for IHD (2,409), ischemic stroke (839), hypertensive heart disease (379) and higher disease burden for alcohol-related diseases, including alcohol use disorders (489) and alcoholic cardiomyopathy (107,58). Cluster 3 (n = 5), the metabolic, perinatal and infectious diseases cluster comprised Central and Latin America countries and Türkiye. This cluster showed the greatest ASDR medians for T2DM (1,013), congenital disorders including neonatal preterm birth (664) and congenital heart anomalies (399), and lower respiratory infections (527,56).

Figure 2. Distribution of ASDR by epidemiological cluster in OECD Countries



Variable-importance analysis using RF algorithms showed that IHD, neonatal preterm birth, Alzheimer's disease and alcohol-use disorders were among the most discriminant conditions across clusters (Supplementary Table 2).

3.2. The Gini index and the burden of disease

The definition and descriptive statistics of the variables used are shown in Table 1. This table reports the pooled summary statistics across the 418 country-year observations from 2009 to 2019. From the studied diseases, the average ASDR per 100,000 was the highest for IHD (mean = 1,549.87, SD = 953.22), followed by anxiety disorders (mean = 585, SD = 162), and neonatal

preterm birth (mean = 199.38). The Gini index had a mean of 34.12 ranging from 23.22 to 54.53 across all country-year observations. Considerable variation was also observed in health system indicators such as healthcare expenditure per capita as percentage of GDP (mean = 8.71%, SD = 2.22) and bed density per 1,000 (mean = 4.59, SD = 2.55). The macroeconomic variable unemployment rates, measured as the percentage of unemployed labour force, had a mean of 6.98% (SD = 4.18). The distribution of risk factors also revealed substantial heterogeneity across observations, for instance, hypertension prevalence ranged from 21.90% in Switzerland in 2019 to 51.20% in Lithuania in 2009, and obesity prevalence ranged from 3.5 in Japan in 2009 to 40.5 in the United States (US) in 2019.

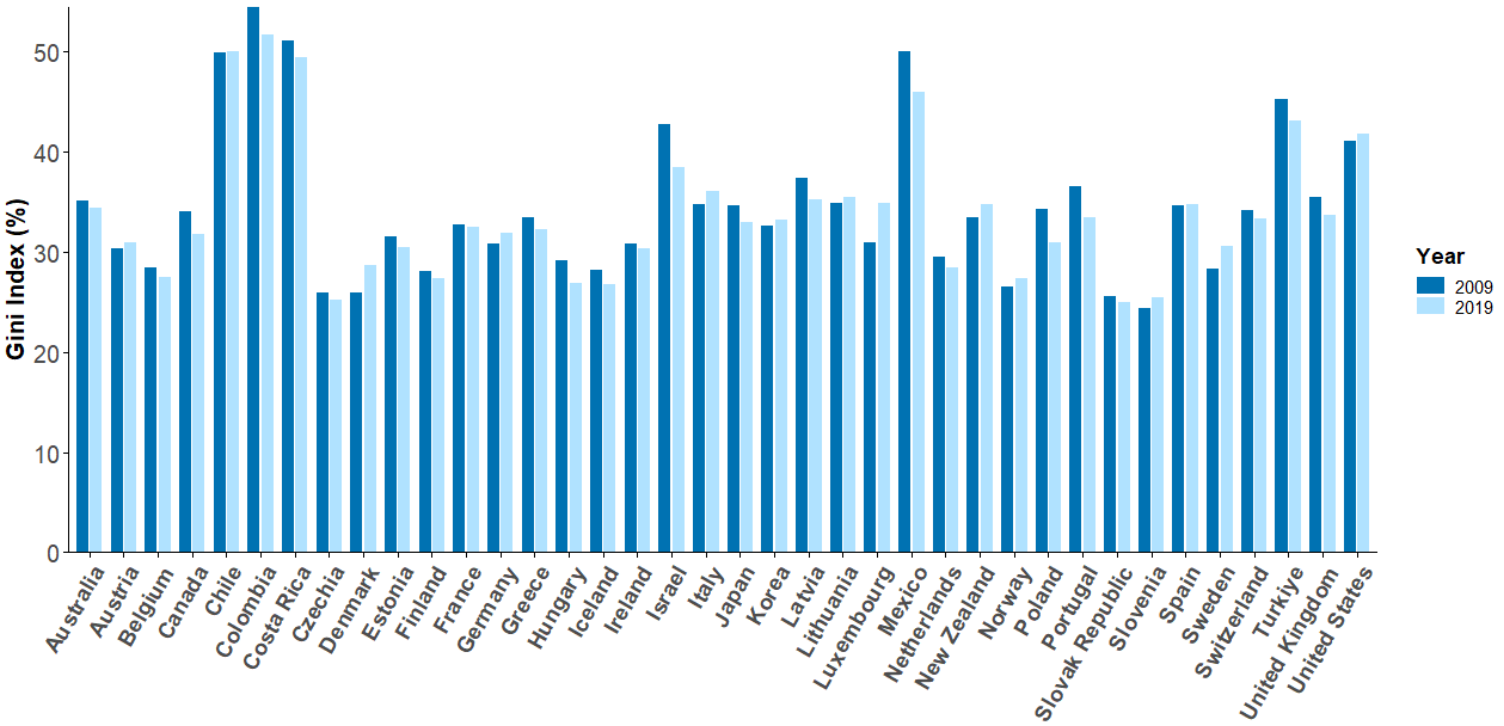
Table 1. Descriptive statistics of the variables studied in OECD countries for the period 2009-2019

Variable	Definition of the variable (Unit)	Observations	Mean	SD	Min	Max
Panel A: Dependent variables						
Anxiety disorders	Age-standardized DALY rate (1/100,000)	418	585	162	284	1051
Ischaemic heart disease	Age-standardized DALY rate (1/100,000)	418	1,549.87	953.22	475.53	4,954.75
Neonatal preterm birth	Age-standardized DALY rate (1/100,000)	418	351.92	199.38	127.74	1,342.57
Panel B: Explanatory variable						
Gini index	0 indicating perfect equality and 100 perfect inequality	418	34.12	6.85	23.22	54.53
Panel C: Control variables						
Healthcare expenditure	% of the GDP per capita	418	8.71	2.22	4.12	16.69
Beds	Number of beds per (1/1,000 inhabitants)	418	4.59	2.55	0.95	13.62
Physicians	Number of physicians (1/1,000 inhabitants)	418	3.31	0.93	1.03	6.25
Public health coverage	Proportion of expenditure on health from the government or compulsory schemes (%)	418	73.35	9.16	46.67	85.75
Unemployment rate	Proportion of the labour force aged 25–64 years that is unemployed (%)	418	6.98	4.18	1.86	28.26

Panel D: Risk factors						
Hypertension	Prevalence of hypertension among adults aged 30-79 years, age-standardized (%)	418	35.17	6.43	21.90	51.20
Physical inactivity	Prevalence of physical inactivity activity among adults aged 18+ years, age-standardized (%)	418	27.53	9.57	10.50	53.30
Tobacco	Proportion of population who are daily smokers in people over 15 years, age-standardized (%)	418	18.77	5.18	6.40	31.90
Obesity	Prevalence of obesity	418	20.7	7.09	3.5	40.5

Income inequality, captured by the Gini index, decreased modestly in 23 OECD countries from 2009 to 2019 (Figure 2, Supplementary Figure 2-4). The greatest declines in Gini index from 2009 to 2019 were found in Israel (-4.33 points in the Gini index), Mexico (-4.01) and Poland (-3.31), whereas income inequality gaps widened in Luxembourg (+3.85), Denmark (+2.75) and Sweden (+2.29, Supplementary Table 5).

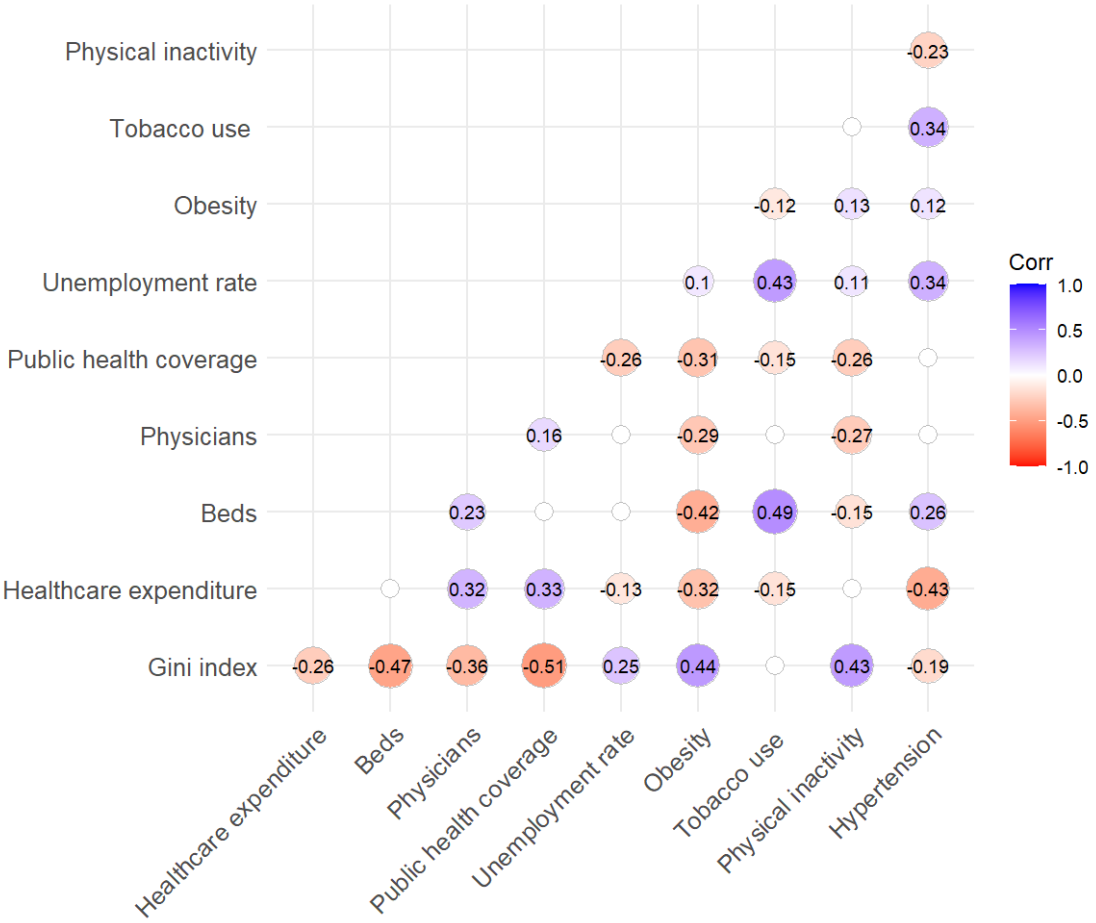
Figure 2. Gini index by country in 2009 and in 2019



Health system variables and macroeconomic indicators evolved in different directions across the OECD during the period of interest. Specifically, healthcare spending as a share of GDP rose in four-fifths of the countries, with the largest increase in Korea (+2.23%) and Japan (+2.01%), while other countries experienced contractions, particularly in Southern and Eastern Europe. Regarding health workforce variables, the number of hospital beds decreased over time, whereas physician density rose in 31 of 35 countries, being accentuated in the Nordic countries. The unemployment rate decreased in most of the countries followed by the post-crisis recovery in 2009 and 2010 (Supplementary Table 5).

The Gini Index was positively correlated with the unemployment rate (Spearman $\rho = 0.25$, $p < 0.001$), while inversely correlated to healthcare expenditure, number of hospital beds, physicians and public health coverage, suggesting that countries with greater expenditure in healthcare and healthcare workforce are also characterized by lower levels of income inequality (Figure 3 and Supplementary Table 6).

Figure 3. Spearman correlations between the studied covariates[4]



[4] Only statistically significant correlations ($p < 0.05$) are shown in the figure. For exact p-values refer to Supplementary Table 6.

Table 2 reports two-way fixed-effects estimates of the within-country association between income inequality and ASDR for IHD, neonatal preterm birth, and anxiety disorders. The direction and magnitude of this association differed by condition. In the models that controlled for health system covariates but not for behavioural and metabolic risk factors, a one-point rise in the Gini index was associated with a 1.5% rise in neonatal preterm birth ASDR (Model 3), a 1.3% decline in IHD ASDR (Model 1) and it was not significant for anxiety disorder (Model 5).

Crude models including only Gini index and country and year fixed effects (Supplementary Table 7) showed the same patterns with negative association ($\beta = -0.012$, SE = 0.006, $p < 0.001$) for IHD, a positive association for neonatal preterm birth ($\beta = 0.018$, SE = 0.003, $p < 0.001$) and no significant association for anxiety disorders. Introducing health system variables (Models 1,3,5) slightly decreased the neonatal coefficient from 0.018 to 0.015, but left the direction and order of magnitude of the IHD coefficient almost unchanged, from -0.012 to -0.013.

Among the health-system covariates, higher physician density per 1000 and greater public health coverage were associated with higher IHD burden, whereas healthcare expenditure was inversely associated to IHD ASDR. However, the association with physician density lost significance after adding risk factors. For neonatal preterm birth, higher government health expenditure and greater number of beds per 1000 were significantly associated with lower ASDR, while public health coverage and unemployment rates were linked to a greater burden. None of these covariates remained significant after conditioning on risk factors. For anxiety disorders, public health coverage was negatively associated with disability rates whereas unemployment rates were positively linked to it. Both effects persisted in the fully adjusted model and physician density also became negatively associated with disease burden.

To examine whether modifiable risk factors help account for the association between income inequality and ASDR burden, smoking, physical inactivity, hypertension and obesity were included in the analysis (Models 2, 4 and 6). The Gini coefficient remained significant for two diseases, IHD ($\beta = -0.012$, $p<0.01$) and neonatal preterm birth ($\beta = 0.011$, $p<0.05$), indicating a persistent within-country association beyond those risks. Tobacco prevalence was positively correlated with IHD burden while for neonatal preterm birth, none of the risk factors were associated with the outcome. When considering anxiety disorders, hypertension and obesity prevalence were positively associated with anxiety disability rates whereas physical inactivity showed the contrary effect. Overall, while behavioural risk factors accounted for a portion of the variance in ASDR, income inequality remained independently associated with IHD and neonatal preterm birth ASDR.

Table 2. Country and year fixed effect model with the covariates

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	IHD	IHD and risk factors	Neonatal preterm birth	Neonatal preterm birth and risk factors	Anxiety disease	Anxiety disease and risk factors
Gini index	-0.013 (0.005)*	-0.012 (0.004)**	0.015 (0.002)***	0.011 (0.005)*	-0.001 (0.002)	-0.0001 (0.001)
Healthcare expenditure	-0.027 (0.013)*	-0.021 (0.010)*	-0.014 (0.007)*	-0.017 (0.015)	0.005 (0.005)	0.007(0.001)
Beds	0.032 (0.019)	0.019 (0.022)	-0.020 (0.009)*	0.001 (0.019)	-0.004 (0.003)	0.0004 (0.003)
Physicians	0.074 (0.028)**	0.032 (0.028)	0.015 (0.010)	0.039 (0.037)	-0.019 (0.021)	-0.014 (0.006)*
Public health coverage	0.004 (0.001)***	0.003 (0.001)***	0.002 (0.001)**	0.003 (0.002)	-0.001 (0.000)*	-0.001 (0.001)*
Unemployment rate	0.003 (0.003)	0.001 (0.002)	0.007 (0.002)**	0.007 (0.004)	0.002 (0.001)*	0.002 (0.001)**
Hypertension		-0.022 (0.011)		-0.004 (0.015)		0.005 (0.002)*
Physical inactivity		0.006 (0.005)		-0.007 (0.006)		-0.003 (0.0005)***
Tobacco		0.011 (0.003)**		-0.008 (0.003)		-0.0002 (0.0005)
Obesity		0.023 (0.013)		-0.005 (0.011)		0.004 (0.001)**
R-squared	0.085	0.365	0.132	0.189	0.125	0.181
Hausman test P-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Note: The cluster robust standard errors are used in the estimation[5]. *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$.

[5] Since the observations of the same country in different years presented serial autocorrelation, and the assumption of standard error requires that the standard error is independent and identically distributed, the estimation of common standard error is not accurate, so this study reports the cluster robust standard error.

4. Discussion

This study provides a clustering analysis of OECD countries based on their epidemiological profiles and a detailed examination of the association between income inequality and ASDR for three diseases, neonatal preterm birth, anxiety disorders and IHD. The findings indicate that the association between the Gini index and disease burden is not uniform across the studied conditions.

The clustering analysis grouped OECD countries into three clusters based on their ASDR for their top 10 diseases. The first cluster included Western Europe, the US, Commonwealth countries, Japan and Korea, and was characterized by high disability rates of mental health and neurodegenerative diseases. The second cluster grouping Eastern European countries was described to have a high ASDR for cardiovascular and alcoholic-related diseases. Finally, the third cluster comprising Central and Latin America countries and Türkiye showed a high burden of metabolic, perinatal and infectious diseases. This is the first study to cluster the 38 OECD countries based on their ASDR. Previous studies have used this unsupervised method to group world countries based on their healthy life expectancy (51), DALYs and risk factors (46), and specifically for OECD countries, as a tool to examine how countries share similar health system characteristics and its effect on health performance (45). However, our findings are in line with the ones described by the Global Burden of Disease 2016 study reporting that countries with middle socio-demographic development index (SDI) present a higher DALYs burden in communicable, maternal, neonatal and nutritional diseases compared to high SDI countries, where non-communicable diseases are the main cause of disability (52).

We have shown that higher income inequality is linked to higher ASDR of neonatal preterm birth, even after accounting for health system characteristics and country-level prevalence of behavioural and metabolic risk factors. Previous investigations of the association between income inequality and preterm birth were single-country and included individual-level which makes them not directly comparable to this ecological study. However, their findings go in the same direction as our results, such as the study of Nkansah-Amankra et al. (2010) which found that mothers living in medium and high Gini neighbourhoods at the time of birth were at increased risk of preterm births compared to mothers living in low-income inequality neighbourhoods (53). In addition, Huynh et al. (2005) reported that pregnant women residing in areas with medium and high-income inequality experienced 7% and 23% increased risks of preterm birth, respectively, compared with those in low-inequality regions (54). Furthermore, the idea that increased prenatal preterm birth disability burden correlates with higher income inequality could be explained by resource deprivation linked with the neo-materialistic theory. This theory suggests that a systematic lack of investment in public services could disproportionately affect the poorer segment of the population who face difficulties in preventing

and treating disease (55). Consistent with this theory, and without adjusting for risk factors, greater healthcare expenditure and bed density were associated with lower preterm birth ASDR while the national unemployment rate was showing the opposite direction. At the individual level, a cross-sectional study in France showed that women who were not working during pregnancy had less adequate preventive behaviour and a higher risk of adverse perinatal outcomes than pregnant women working in high or intermediate professional categories (56). Regarding government or compulsory schemes as a percentage of expenditure on health, higher public health coverage showed a positive linked with a higher burden of prenatal preterm births. This could be partially explained by the context of the US, where public health insurance programs such as Medicaid and Medicare often serve as proxies of socioeconomic disadvantage. A cohort study conducted in 11 US states on 223,512 preterm birth children revealed that both income inequality and public insurance were associated with higher rates of preterm birth (57). This exposes a case of selection bias, as individuals covered by public insurance are more likely to belong to vulnerable populations at elevated risk for adverse perinatal outcomes. Country-level risk factors were not associated with preterm birth DALY rates, but when they were adjusted for, only the Gini coefficient remained significant, although its effect size was attenuated. This may suggest these risk factors may help explain part of the observed variation in the association between income inequality and neonatal preterm birth, although a direct mediating role cannot be confirmed in this study. In addition, since these factors are measured at the country level rather than the individual level, inferring at the individual-level would fall into an ecological fallacy.

In the analysis of IHD, we report that higher income inequality is associated with lower DALY rates, an effect that remains even after adjusting for risk factors. Previous cross-sectional studies demonstrate partial agreement with the present thesis' findings. An investigation of 16 Luxembourg Inequality Survey countries, including only high-income countries, reported non-significant correlations between the Gini coefficient and coronary heart disease mortality rates adjusting for GDP per capita, although the authors used mortality as their dependent variable (58). Furthermore, another study demonstrated inverse associations between income inequality and age-specific cardiovascular mortality rates in women over 50 years old. However, the main body of literature shows positive linkages between income inequality and both IHD morbidity and mortality in OECD countries (59), as well as worse outcomes in heart failure in larger samples (60).

Several factors may help explain the discrepancies between our findings and those of previous studies reporting a positive association between income inequality and IHD burden. First, methodological differences play an important role. This ecological study uses country-level data and two-way fixed effects, which remove all time-invariant cultural, dietary and climatic traits. As

a result, cross-country differences that may influence cardiovascular outcomes are not captured nor specific regional contexts within countries. Moreover, individual-level factors, such as personal income, healthcare access or health behaviours were not included. This limitation could explain the difference with the findings observed in several cohorts at the individual level, which consistently show that lower socioeconomic status is associated with worse cardiovascular health outcomes (61–63). Second, in the studied period from 2009 to 2019, most OECD countries experienced significant improvements in cardiovascular prevention and treatment strategies. These could be due to the reduction in smoking prevalence, which over the last decade has decreased in the majority the OECD countries (64), an improved hypertension control, detection and treatment in high-income countries from 1990 to 2019 (65) and the implementation of trans-fat bans, which were associated with significant reductions in trans fatty acids levels (66). Moreover, another major contributor to the reduction of IHD burden was the widespread use of statins over the past two decades for primary and secondary prevention of cardiovascular disease which have substantially reduced the risk of major coronary events (67). Finally, the adoption of primary percutaneous coronary intervention for ST-elevation myocardial infarction that has reduced long-term morbidity and improved survival in patients with acute coronary syndromes (68). Taken this together, these public health advancements have likely contributed to a decline in the overall IHD burden across OECD countries. However, these benefits have not been distributed equally across the different socioeconomic groups as illustrated by previous studies. As Clark et al. (2009) reports, patients with lower socioeconomic status are less likely to receive evidence-based pharmacological treatments, such as statins, cardiac intervention including primary percutaneous coronary intervention, behavioural therapies or follow up with a cardiologist (69). In addition, while individuals from lower socioeconomic groups have higher exposure to modifiable risk factors such as smoking (70), they are also the ones that are less likely to quit smoking, even after cessation programs (71). Furthermore, socioeconomic inequalities are associated to a worse control of hypertension (72) and access to cardiac interventions (73). These disparities may mean that, even though national rates of IHD decline, inequalities within countries could remain hidden or worsen. Therefore, countries with rising income inequalities could be reporting improvements in cardiovascular health that would be just benefiting the more advantaged groups. Third, the crude model including only the Gini index and IHD DALYs also reported a negative association between both of them, suggesting that this effect persists even after adjusting for health system confounders and health risk factor prevalence variables. Fourth, the Gini index in our study ranged from 0.23 to 0.55[6], which is narrower and lower compared to the global income inequality distribution. For instance, a longitudinal study of income inequality and

[6] In the present study, the Gini index is expressed on a scale from 1 to 100; however, it can also be presented on a scale from 0 to 1.

its impact on non-communicable disease burden in different Brazilian states showed a Gini index range from 0.66 to 0.89 (43). This could explain the differences between our results and other global cross-country studies.

Regarding health system financing, higher healthcare expenditure per capita was linked to reduced IHD disability rates, whereas a greater share of public health coverage showed the opposite pattern, mirroring our prenatal preterm birth findings. Healthcare system resources, proxied by the number of physicians, were associated with a higher burden of IHD ASDR. It is possible that countries with higher IHD ASDR, *ceteris paribus*, spend more on health and employ more physicians per head of population to address this healthcare need. IHD is a multifactorial condition driven by a complex interplay of environmental, metabolic and lifestyle exposures. Specifically for modifiable risk factors, a GBD analysis covering 204 countries from 1990 to 2019 revealed that hypertension, high low-density lipoprotein cholesterol and smoking were the main contributors to IHD DALY burden (74). In our study, only the prevalence of tobacco use remained significantly associated with higher IHD disability rates, whereas other GBD risk factors did not show independent effects. After adjustment for risk factors, the Gini coefficient, healthcare expenditure per capita and public scheme coverage retained their significant associations with IHD ASDR. Notably, the inverse association between inequality and IHD burden was almost not changed, decreasing only slightly in magnitude from -0.013 to -0.012 and retaining statistical significance. The persistence of a significant Gini effect highlights that the association between income inequality and IHD burden is not fully explained by the included risk factors.

In the case of anxiety, income inequality was not associated with ASDR. However, greater unemployment rates were associated with a higher anxiety burden, both before and after including risk factors, which is consistent with previous literature (75,76), including an investigation across 201 countries from 1970 to 2020 which revealed that a 1% increase in unemployment rates corresponds to a 0.0087% increase in anxiety disorders prevalence (77). On the contrary, physicians' number and public health coverage were associated with lower ASDR for anxiety, the first being significant only when conditioning for risk factors. Assuming that greater physician density reflects increased availability of mental-health specialists, this likely enhances access to diagnostic and therapeutic services, thereby attenuating the population burden of anxiety. Similarly, broader public scheme coverage associated with reduced anxiety disability, suggests that universal health coverage frameworks which minimize out-of-pocket costs for mental health services are linked to better mental health outcomes and a lower prevalence of untreated anxiety disorder. Regarding risk factors, hypertension and obesity were positively associated with anxiety ASDR, whereas physical inactivity showed the opposite effect. For obesity, these findings are supported by previous studies, such as the cross-

sectional population-based survey study on 62,277 adults that demonstrated a significant pooled odds ratio of 1.2 for total obesity and 1.4 for severe obesity with anxiety disorders (78) and others (79,80). Regarding hypertension, a positive link with anxiety has been previously reported in previous literature, including an exhaustive systematic review of cross-sectional and prospective studies (81). The negative association between physical inactivity and anxiety disability rates is unexpected, as the main body reports that sedentary behaviours are a risk for anxiety (82,83). However, this ecological study does not capture individual-level behaviours which could be influencing anxiety burden.

Our analysis has limitations. First, the unit of analysis was the country level which could have masked more detailed associations between the Gini index and health at the individual level. Second, and also related to the quality of the databases used, several variables exhibited missing observations that were imputed using forward and backward-filling. Because these imputations rely on adjacent values, they may artificially decrease the true within-country over time variance. Third, its complete reliance on the data quality and processing standards of the secondary databases used could have introduced consistencies. This is mainly the case for the GBD database from which ASDR were obtained for each country and year. Although the IHME, in collaboration with the WHO, provides regularly updated estimates, these are modelled rather than directly observed data and may diverge from national statistics. Moreover, GBD does not capture the burden of small population subgroups, including rare and ultra-rare diseases (84). Nonetheless, despite these limitations, the GBD remains the only global epidemiological resource which offers comparable estimates across OECD countries. Fourth, the use of a two-way fixed effect over a random effect model, recommended by the Hausman test, captured only the within-country variation but ignored the between-country differences, which were also informative. We acknowledge that the models R^2 are modest, however, as Deaton (2003) stated, studies that apply fixed-effects models tend to find weaker associations between inequality and health, possibly because they account for unobserved, time-invariant country characteristics that confound simpler cross-sectional relationships (85). Fifth, to avoid collinearity with the Gini index, we excluded health financing and sociodemographic variables commonly used in inequality research such as GDP per capita and education, which could be confounders of the association between health inequality and the burden of disease. Sixth, we acknowledge that applying the same set of risk factors across different diseases, despite their unique aetiologies, could miss disease-specific pathways and introduce residual confounding into our estimates. Seventh, the analysis did not account for gender specific effects, which could have been important given the evidence that inequality affects men and women differently depending on disease type (86). Finally, causal interpretations cannot be made due to the ecological design of this study. Therefore, further research, ideally targeted to specific populations and periods, is essential to elucidate the true causal impact of income inequality on health outcomes.

Taken together, our findings reinforce the view that macro-level economic variables set the parameters within which individual risk unfolds. Across three distinct conditions, neonatal pre-term birth, IHD and anxiety disorders, we observed that indicators of income distribution, public health coverage and the design of health-financing systems maintained their association with the burden of disease even after controlling for the main risk factors. In this context, advocating exclusively for a behavioural explanation of health inequalities can lead to the growth of "healthism", an ideology that promotes the idea that maintaining good health is a matter of personal choice and discipline, and where the individual is the ultimate responsible for improving their lifestyle (87). It would be misleading to separate behaviour and lifestyle from the social structure, since the ability to make healthy choices is shaped by socioeconomic constraints and the availability of options (9).

Policy implications in the macro-level context follow directly. If OECD governments wish to achieve the Sustainable Development Goal targets on maternal, cardiovascular and mental health, interventions must extend beyond individual behavioural interventions. They should also consider redistributive labour-market, fiscal and welfare policies that narrow the income gap and guarantee equitable access to high-quality care throughout the life course. In addition, improving data infrastructures is essential for correctly monitoring disease burden. Public health surveillance systems should integrate regional data on income inequality, health system performance and health access to better capture health differences across population groups. Finally, stronger collaborations between public health, education, housing and employment sectors is needed to address the social determinants of health more holistically. Without tackling the upstream determinants of health, such as income inequalities, downstream interventions are likely to have limited effectiveness, particularly for complex and chronic conditions including IHD and anxiety.

> "*Neoliberals – New Labor for example – can appear quite progressive about gender, race, sexuality, disability and condemn those who discriminate against people on these grounds. Unsurprisingly, the elephant in the room is economic inequalities or class difference."* (88)

5. Conclusion

The three clusters obtained through hierarchical clustering of OECD countries reveal different epidemiological profiles, with a notable regional distribution pattern. Across the selected representative diseases, income inequality was positively associated neonatal preterm birth ASDR but negatively linked with IHD ASDR, a finding that warrants cautious interpretation. At the macro level, greater healthcare expenditure per capita and lower unemployment rates were consistently associated with lower disease burden. After adjusting for behavioural and metabolic risk factors, the association between income inequality and disease burden was slightly attenuated suggesting that these risk factors may explain a small part of the observed variation. However, the persistence of the association between income inequality and health persists after adjustment indicates that policies limited to behaviour change are unlikely to be sufficient to reduce health inequalities. Instead, equity-oriented labour-market policies, enhanced social protection and sustained investment in healthcare should also be considered. The findings highlight the importance of tackling the broader social determinants of health to improve population well-being. Finally, the cluster-specific epidemiological profile identified in this study may help guide public health interventions tailored to the specific health needs of OECD countries.

6. References

1. Zhavoronkov A, Debonneuil E, Mirza N, Artyuhov I. Evaluating the impact of recent advances in biomedical sciences and the possible mortality decreases on the future of health care and Social Security in the United States. Pensions Int J. 2012 Nov;17(4):241–51.

2. Boudoulas KD, Triposkiadis F, Stefanadis C, Boudoulas H. The endlessness evolution of medicine, continuous increase in life expectancy and constant role of the physician. Hellenic J Cardiol. 2017 Sep;58(5):322–30.

3. Gregory IN. Comparisons between geographies of mortality and deprivation from the 1900s and 2001: spatial analysis of census and mortality statistics. BMJ. 2009 Sep 11;339(sep10 2):b3454–b3454.

4. Educational inequalities in longevity among OECD countries around 2016 [Internet]. 2022 Aug [cited 2025 Jun 3]. (OECD Papers on Well-being and Inequalities; vol. 8). Report No.: 8. Available from: https://www.oecd.org/en/publications/educational-inequalities-in-longevity-among-oecd-countries-around-2016_5faaa751-en.html

5. Whitehead M. Life and death over the millennium. In: Health inequalities. Lodon: The Stationery Office; 1997.

6. Taylor R, Rieger A. Rudolf Virchow on the typhus epidemic in Upper Silesia: an introduction and translation. Sociol Health Illn. 1984 Jul;6(2):201–17.

7. Marmot MG, Stansfeld S, Patel C, North F, Head J, White I, et al. Health inequalities among British civil servants: the Whitehall II study. The Lancet. 1991 Jun;337(8754):1387–93.

8. Dahlgren D, Whitehead M. Policies and Strategies to Produce Social Equity in Health. Stockholm: Institute for Futures Studies; 1991.

9. Wilkinson RG, Marmot MG, Weltgesundheitsorganisation, editors. Social determinants of health: the solid facts. 2nd. ed. Copenhagen: WHO Regional Office for Europe; 2003. 31 p.

10. Dumas A, Robitaille J, Jette SL. Lifestyle as a choice of necessity: Young women, health and obesity. Soc Theory Health. 2014 May;12(2):138–58.

11. Ogunlayi F, Coleman PC, Fat LN, Mindell JS, Oyebode O. Trends in socioeconomic inequalities in behavioural non-communicable disease risk factors: analysis of repeated cross-sectional health surveys in England between 2003 and 2019. BMC Public Health. 2023 Jul 28;23(1):1442.

12. Allen L, Williams J, Townsend N, Mikkelsen B, Roberts N, Foster C, et al. Socioeconomic status and non-communicable disease behavioural risk factors in low-income and lower-middle-income countries: a systematic review. Lancet Glob Health. 2017 Mar;5(3):e277–89.

13. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health. 2020 Nov;74(11):964–8.

14. United Nations. Goal 10: Reduce inequality within and among countries [Internet]. Available from: https://www.un.org/sustainabledevelopment/inequality/

15. Lynch J, Smith GD, Harper S, Hillemeier M, Ross N, Kaplan GA, et al. Is Income Inequality a Determinant of Population Health? Part 1. A Systematic Review. Milbank Q. 2004 Mar;82(1):5–99.

16. Marmot M. Psychosocial and material pathways in the relation between income and health: a response to Lynch et al. BMJ. 2001 May 19;322(7296):1233–6.

17. Putnam R. The Prosperous Community: Social Capital and Public Life. Vol. 13. The American Prospect; 1993. 35–42 p.

18. Yao Y, Wan G, Meng D. Income distribution and health: can polarization explain health outcomes better than inequality? Eur J Health Econ. 2019 Jun;20(4):543–57.

19. Luptáčik M, Nežinský E. Measuring income inequalities beyond the Gini coefficient. Cent Eur J Oper Res. 2020 Jun;28(2):561–78.

20. Subramanian SV. Income Inequality and Health: What Have We Learned So Far? Epidemiol Rev. 2004 Jul 1;26(1):78–91.

21. Fiscella K, Franks P. Poverty or income inequality as predictor of mortality: longitudinal cohort study. BMJ. 1997 Jun 14;314(7096):1724–1724.

22. Wagstaff A, Van Doorslaer E. Income Inequality and Health: What Does the Literature Tell Us? Annu Rev Public Health. 2000 May;21(1):543–67.

23. Judge K. Income distribution and life expectancy: a critical appraisal. BMJ. 1995 Nov 11;311(7015):1282–5.

24. Varbanova V, Hens N, Beutels P. Determinants of life-expectancy and disability-adjusted life years (DALYs) in European and Organisation for Economic Co-operation and Development (OECD) countries: A longitudinal analysis (1990–2019). SSM - Popul Health. 2023 Dec;24:101484.

25. Irandoust K, Daroudi R, Tajvar M, Yaseri M. Assessing health determinants worldwide: Econometric analysis of the Global Burden of Diseases Study 2000–18 – Highlighting impactful factors on DALY, YLL, and YLD indicators. J Glob Health. 2024 Mar 15;14:04051.

26. Shi L, Macinko J, Starfield B, Wulu J, Regan J, Politzer R. The Relationship Between Primary Care, Income Inequality, and Mortality in US States, 1980-1995. J Am Board Fam Med. 2003 Sep 1;16(5):412–22.

27. Park J, Ryu SY, Han M ah, Choi SW. The association between income inequality and all-cause mortality across urban communities in Korea. BMC Public Health. 2015 Dec;15(1):574.

28. Materia E, Cacciani L, Bugarini G, Cesaroni G, Davoli M, Mirale MP, et al. Income inequality and mortality in Italy. Eur J Public Health. 2005 Aug 1;15(4):411–7.

29. Naghavi M, Ong KL, Aali A, Ababneh HS, Abate YH, Abbafati C, et al. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet. 2024 May;403(10440):2100–32.

30. Ferrari AJ, Santomauro DF, Aali A, Abate YH, Abbafati C, Abbastabar H, et al. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries

and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet. 2024 May;403(10440):2133–61.

31. VizHub Global Burden of Disease results. Available from: https://vizhub.healthdata.org/gbd-results/

32. Global Health Obsevatory by WHO. Available from: https://www.who.int/data/gho

33. OECD Data Explorer. Available from: https://data-explorer.oecd.org/

34. World Bank Open Data. Available from: https://data.worldbank.org/

35. World Income Inequality Database. Available from: https://www4.wider.unu.edu/?ind=40&type=ChoroplethSeq&year=70&byCountry=false&slider=buttons

36. Maher C, Ferreira G. Time to reconsider what Global Burden of Disease studies really tell us about low back pain. Ann Rheum Dis. 2022 Mar;81(3):306–8.

37. Charles V, Gherman T, Paliza JC. The Gini Index: A Modern Measure of Inequality. In: Charles V, Emrouznejad A, editors. Modern Indices for International Economic Diplomacy [Internet]. Cham: Springer International Publishing; 2022 [cited 2025 May 26]. p. 55–84. Available from: https://link.springer.com/10.1007/978-3-030-84535-3_3

38. Gini C. Measurement of Inequality of Incomes. Econ J. 1921 Mar;31(121):124.

39. OECD. Rethinking Health System Performance Assessment: A Renewed Framework [Internet]. OECD; 2024 [cited 2025 May 26]. (OECD Health Policy Studies). Available from: https://www.oecd.org/en/publications/rethinking-health-system-performance-assessment_107182c8-en.html

40. Rocco L, Fumagalli E, Mirelman AJ, Suhrcke M. Mortality, morbidity and economic growth. Montefiori M, editor. PLOS ONE. 2021 May 27;16(5):e0251424.

41. Berger MC, Messer J. Public financing of health expenditures, insurance, and health outcomes. Appl Econ. 2002 Nov;34(17):2105–13.

42. Brauer M, Roth GA, Aravkin AY, Zheng P, Abate KH, Abate YH, et al. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet. 2024 May;403(10440):2162–203.

43. Gaspar RS, Rossi L, Hone T, Dornelles AZ. Income inequality and non-communicable disease mortality and morbidity in Brazil States: a longitudinal analysis 2002-2017. Lancet Reg Health - Am. 2021 Oct;2:100042.

44. Raghupathi V, Raghupathi W. The influence of education on health: an empirical assessment of OECD countries for the period 1995–2015. Arch Public Health. 2020 Dec;78(1):20.

45. OECD, The Health Foundation. How Do Health System Features Influence Health System Performance? [Internet]. OECD Publishing; 2025 [cited 2025 May 27]. Available from: https://www.oecd.org/en/publications/how-do-health-system-features-influence-health-system-performance_7b877762-en.html

46. Yilmaz Işikhan S, Güleç D. The Clustering of World Countries Regarding Causes of Death and Health Risk Factors. Iran J Public Health. 2018 Oct;47(10):1520–8.

47. Hong Han, Xiaoling Guo, Hua Yu. Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. In: 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS) [Internet]. Beijing, China: IEEE; 2016 [cited 2025 May 27]. p. 219–24. Available from: http://ieeexplore.ieee.org/document/7883053/

48. Izadi N, Yari-Boroujeni R, Soofi M, Niknam M, Amiri P, Najafi F. Socioeconomic inequalities and dyslipidaemia in adult population of the Ravansar Non-Communicable Disease Cohort Study: the role of sex and age. BMJ Open. 2024 Nov;14(11):e085035.

49. Luo T, Lin S, Zhang W, Li X, Wang Y, Zhou J, et al. Relationship between socioeconomic status and hypertension incidence among adults in southwest China: a population-based cohort study. BMC Public Health. 2024 May 2;24(1):1211.

50. United Nations. World Income Inequality Database. Available from: https://www.wider.unu.edu/project/world-income-inequality-database-wiid

51. Levantesi S, Nigri A, Piscopo G, Spelta A. Multi-country clustering-based forecasting of healthy life expectancy. Qual Quant. 2023 Dec;57(S2):189–215.

52. Hay SI, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017 Sep;390(10100):1260–344.

53. Nkansah-Amankra S, Dhawain A, Hussey JR, Luchok KJ. Maternal Social Support and Neighborhood Income Inequality as Predictors of Low Birth Weight and Preterm Birth Outcome Disparities: Analysis of South Carolina Pregnancy Risk Assessment and Monitoring System Survey, 2000–2003. Matern Child Health J. 2010 Sep;14(5):774–85.

54. Huynh M, Parker JD, Harper S, Pamuk E, Schoendorf KC. Contextual effect of income inequality on birth outcomes. Int J Epidemiol. 2005 Aug 1;34(4):888–95.

55. Lynch JW. Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. BMJ. 2000 Apr 29;320(7243):1200–4.

56. Vigoureux S, Lorthe E, Blondel B, Ringa V, Saurel-Cubizolles MJ. Women's occupational status during pregnancy and preventive behaviour and health outcomes between 1998 and 2016 in France. J Gynecol Obstet Hum Reprod. 2023 Mar;52(3):102545.

57. Wallace ME, Mendola P, Chen Z, Hwang BS, Grantz KL. Preterm Birth in the Context of Increasing Income Inequality. Matern Child Health J. 2016 Jan;20(1):164–71.

58. Lynch J, Smith GD, Hillemeier M, Shaw M, Raghunathan T, Kaplan G. Income inequality, the psychosocial environment, and health: comparisons of wealthy nations. The Lancet. 2001 Jul;358(9277):194–200.

59. Kim D, Kawachi I, Hoorn SV, Ezzati M. Is inequality at the heart of it? Cross-country associations of income inequality with cardiovascular diseases and risk factors. Soc Sci Med. 2008 Apr;66(8):1719–32.

60. Dewan P, Rørth R, Jhund PS, Ferreira JP, Zannad F, Shen L, et al. Income Inequality and Outcomes in Heart Failure. JACC Heart Fail. 2019 Apr;7(4):336–46.

61. Zhu Y, Llamosas-Falcón L, Kerr WC, Rehm J, Probst C. Behavioral risk factors and socioeconomic inequalities in ischemic heart disease mortality in the United States: A causal mediation analysis using record linkage data. PLoS Med. 2024 Sep;21(9):e1004455.

62. Zhu Y, Wang Y, Shrikant B, Tse LA, Zhao Y, Liu Z, et al. Socioeconomic disparity in mortality and the burden of cardiovascular disease: analysis of the Prospective Urban Rural Epidemiology (PURE)-China cohort study. Lancet Public Health. 2023 Dec;8(12):e968–77.

63. Mullachery PH, Vela E, Cleries M, Comin-Colet J, Nasir K, Diez Roux AV, et al. Inequalities by Income in the Prevalence of Cardiovascular Disease and Its Risk Factors in the Adult Population of Catalonia. J Am Heart Assoc. 2022 Sep 6;11(17):e026587.

64. OECD. Health at a Glance 2023: OECD Indicators [Internet]. OECD; 2023 [cited 2025 Jun 2]. (Health at a Glance). Available from: https://www.oecd.org/en/publications/health-at-a-glance-2023_7a7afb35-en.html

65. Zhou B, Carrillo-Larco RM, Danaei G, Riley LM, Paciorek CJ, Stevens GA, et al. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. The Lancet. 2021 Sep;398(10304):957–80.

66. Downs SM, Thow AM, Leeder SR. The effectiveness of policies for reducing dietary trans fat: a systematic review of the evidence. Bull World Health Organ. 2013 Apr 1;91(4):262-269H.

67. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. The Lancet. 2016 Nov;388(10059):2532–61.

68. Widimsky P, Wijns W, Fajadet J, De Belder M, Knot J, Aaberge L, et al. Reperfusion therapy for ST elevation acute myocardial infarction in Europe: description of the current situation in 30 countries. Eur Heart J. 2010 Apr 2;31(8):943–57.

69. Clark AM, DesMeules M, Luo W, Duncan AS, Wielgosz A. Socioeconomic status and cardiovascular disease: risks and implications for care. Nat Rev Cardiol. 2009 Nov;6(11):712–22.

70. Stringhini S, Carmeli C, Jokela M, Avendaño M, Muennig P, Guida F, et al. Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. The Lancet. 2017 Mar;389(10075):1229–37.

71. Hiscock R, Dobbie F, Bauld L. Smoking Cessation and Socioeconomic Status: An Update of Existing Evidence from a National Evaluation of English Stop Smoking Services. BioMed Res Int. 2015;2015:1–10.

72. Amini M, Moradinazar M, Rajati F, Soofi M, Sepanlou SG, Poustchi H, et al. Socioeconomic inequalities in prevalence, awareness, treatment and control of hypertension: evidence from the PERSIAN cohort study. BMC Public Health. 2022 Jul 22;22(1):1401.

73. Alter DA. Socioeconomic Status, Service Patterns, and Perceptions of Care Among Survivors of Acute Myocardial Infarction in Canada. JAMA. 2004 Mar 3;291(9):1100.

74. Safiri S, Karamzad N, Singh K, Carson-Chahhoud K, Adams C, Nejadghaderi SA, et al. Burden of ischemic heart disease and its attributable risk factors in 204 countries and territories, 1990–2019. Eur J Prev Cardiol. 2022 Mar 11;29(2):420–31.

75. Konstantakopoulos G, Pikouli K, Ploumpidis D, Bougonikolou E, Kouyanou K, Nystazaki M, et al. The impact of unemployment on mental health examined in a community mental health unit during the recent financial crisis in Greece. Psychiatriki. 2020 Jan 2;30(4):281–90.

76. Flatau P, Galea J, Petridis R. Mental Health and Wellbeing and Unemployment. Aust Econ Rev. 2000 Jun;33(2):161–81.

77. Yang Y, Niu L, Amin S, Yasin I. Unemployment and mental health: a global study of unemployment's influence on diverse mental disorders. Front Public Health. 2024 Dec 13;12:1440403.

78. Scott KM, Bruffaerts R, Simon GE, Alonso J, Angermeyer M, De Girolamo G, et al. Obesity and mental disorders in the general population: results from the world mental health surveys. Int J Obes. 2008 Jan;32(1):192–200.

79. Gariepy G, Nitka D, Schmitz N. The association between obesity and anxiety disorders in the population: a systematic review and meta-analysis. Int J Obes. 2010 Mar;34(3):407–19.

80. Balkoca F, Cubukcu M, ŞiMşek Yurt N. The Relationship Between Obesity with Depression and Anxiety Levels: a Cross-Sectional, Case-Control Study. Turk J Diabetes Obes. 2023 Aug 31;7(2):152–8.

81. Lim LF, Solmi M, Cortese S. Association between anxiety and hypertension in adults: A systematic review and meta-analysis. Neurosci Biobehav Rev. 2021 Dec;131:96–119.

82. Silva LRB, Seguro CS, De Oliveira CGA, Santos POS, De Oliveira JCM, De Souza Filho LFM, et al. Physical Inactivity Is Associated With Increased Levels of Anxiety, Depression, and Stress in Brazilians During the COVID-19 Pandemic: A Cross-Sectional Study. Front Psychiatry. 2020 Nov 17;11:565291.

83. Teychenne M, Costigan SA, Parker K. The association between sedentary behaviour and risk of anxiety: a systematic review. BMC Public Health. 2015 Dec;15(1):513.

84. Maher C, Ferreira G. Time to reconsider what Global Burden of Disease studies really tell us about low back pain. Ann Rheum Dis. 2022 Mar;81(3):306–8.

85. Deaton A. Health, Inequality, and Economic Development. J Econ Lit. 2003 Mar;41(1):113–58.

86. Backhans MC, Burström B, Lindholm L, Månsdotter A. Pioneers and laggards – Is the effect of gender equality on health dependent on context? Soc Sci Med. 2009 Apr;68(8):1388–95.

87. Burrow L. Pedagogizing families through obesity discourse. In: J. Wright and V. Harwood (eds.) Biopolitics and the 'Obesity Epidemic'. In New York: Routledge; p. 127–40.

88. Sayer A. Why We Can't Afford the Rich. 2014;

89. Schumacher AE, Kyu HH, Aali A, Abbafati C, Abbas J, Abbasgholizadeh R, et al. Global age-sex-specific mortality, life expectancy, and population estimates in 204 countries and territories and 811 subnational locations, 1950–2021, and the impact of the COVID-19 pandemic: a comprehensive demographic analysis for the Global Burden of Disease Study 2021. The Lancet. 2024 May;403(10440):1989–2056.

7. Appendix. Supplementary materials

7.1. Supplementary Material 1

The burden of disease measures the overall impact of diseases and injuries on a population by combining data on mortality and morbidity. It is primarily expressed using Disability-Adjusted Life Years (DALYs), which represent the total number of healthy years lost due to illness, disability, or premature death. DALYs are calculated as the sum of:

- Years of Life Lost (YLL): due to early death, is a metric that is computed by multiplying the number of estimated deaths by the predicted life expectancy by age, sex, location, and year.
- Years Lived with Disability (YLD): due to illness or injury, calculated by multiplying disease prevalence by a severity weight.

One DALY equals one year of healthy life lost. For example, in a population of 100,000 (89):

- 500 premature deaths, each 10 years earlier than 71.7 years old (global life expectancy at birth) → YLL = 500 x 11,7 = 5,850
- 10,000 people with a disability, average weight 0.2 → YLD = 10,000 × 0.2 = 2,000
- Total burden = DALYs = 5,850 + 2,000 = 7,850

This metric highlights the gap between current health and ideal health where everyone lives a full, disease-free life.

7.2. Supplementary Table 1. Descriptive information on variables, data sources and missingness percentage used across 38 countries from 2009 to 2019

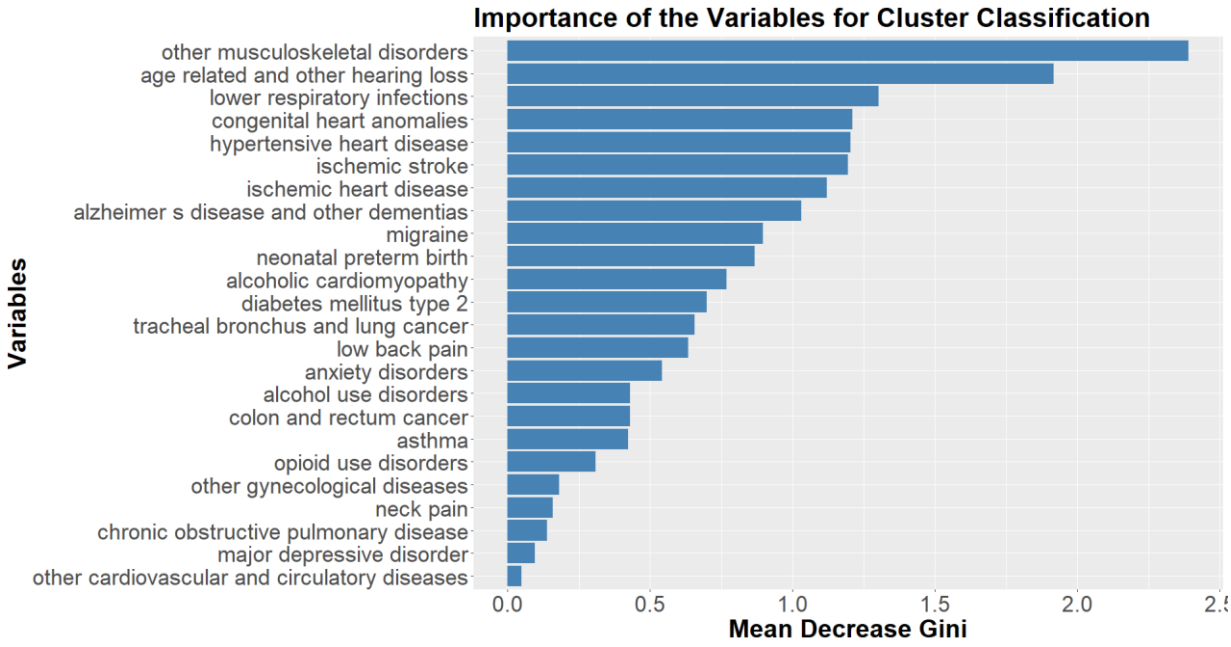
Variable	Unit	Source	Date last day accessed	Missingness percentage
Dependent variable				
Disability-adjusted life years (DALYs)	Age-standardized rate	Global Burden of Disease by Institute for Health Metrics and Evaluation (IHME). Available at: https://vizhub.healthdata.org/gbd-results/?params=gbd-api-2021-public/3add8c44a6706b31633882daf261344c	15/05/2025	0 %
Socioeconomic indicators and care services				
Gini Index	From 0 to 100, 0 indicating perfect equality and 100 perfect inequality	World Income Inequality Database (WIID) by the United Nations. Available at: https://www.wider.unu.edu/project/world-income-inequality-database-wiid	13/05/2025	6.22%.
Palma Ratio	From 0 upwards, with 1 indicating parity between the richest 10% and the poorest 40%.	World Income Inequality Database (WIID) by the United Nations. Available at: https://www.wider.unu.edu/project/world-income-inequality-database-wiid	13/05/2025	0%
Healthcare expenditure	% of the GDP per capita	OECD Data Explorer. Available at: https://data-explorer.oecd.org/fs%5b0%5d=Topic%2C1%7CHealth%25	15/05/2025	0 %

Hospital beds	Beds per 1,000 inhabitants	OECD Data Explorer. Available at: https://data-explorer.oecd.org/?tm=Hospital%20beds%20&pg=0&snb=9	13/05/2025	4.3 %
Physicians rate	Doctors per 1,000 inhabitants	World Data Bank. Available at: https://databank.worldbank.org/reports.aspx?source=2&series=SH.MED.PHYS.ZS&country=JPN,AUS,CAN,FRA,DEU,ESP,ITA,KOR,GBR,USA,MEX	12/05/2025	4.07%
Coverage of public healthcare	Government or compulsory schemes as a percentage of expenditure on health	OECD Data Explorer. Available at: https://data-explorer.oecd.org/vis?tm=health%20insurance&pg=0&hc[Health%20function]=&snb=169&vw=tb&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_SHA%40DF_SHA&df[ag]=OECD.ELS.HD&df[vs]=1.0&dq=USA%2BGBR%2BTUR%2BCHE%2BSWE%2BESP%2BSVK%2BSVN%2BPOL%2BPRT%2BNOR%2BNLD	13/05/2025	0 %
Unemployment rate	% of unemployed labour force from 25 to 64	OECD Data Explorer. Available at: https://data-explorer.oecd.org/vis?tm=DF_IDD&pg=0&snb=1&vw=tb&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_WISE_IDD%40DF_IDD&df[ag]=OECD.WISE.INE&df[vs]=&pd=2010%2C&dq=.A.PR_INC_MRKT..._T.METH2012.D_CUR.&ly[rw]=REF_AREA%2CUNIT_MEASURE&ly[cl]=TIME_PERIOD&to[TI	13/05/2025	1.43%

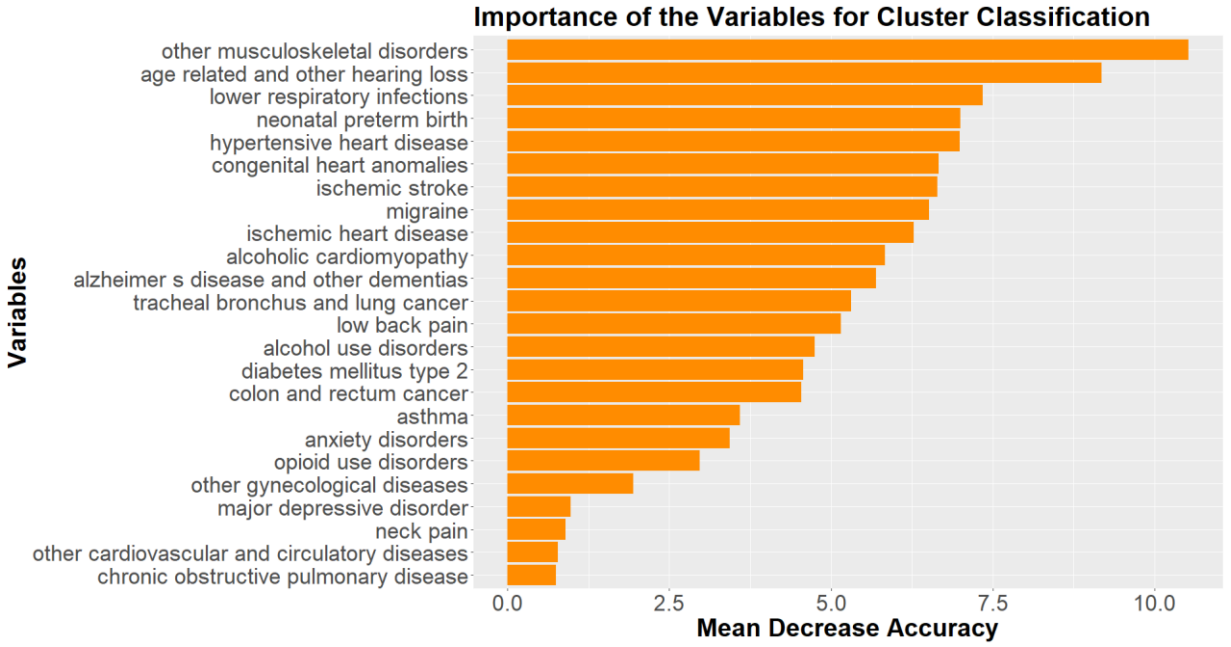
Risk factors				
Hypertension	Prevalence of hypertension among adults aged 30-79 years, age-standardized %	WHO. The Global Health Observatory. Available at: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-hypertension-among-adults-aged-30-79-years	15/05/2025	0%
Obesity	Prevalence of obesity among adults over 18 with a BMI >= 30, age standardized estimate %	WHO. The Global Health Observatory. Available at: *https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-obesity-among-adults-bmi--30-(age-standardized-estimate)-(-)*	15/05/2025	0%
Insufficient physical	Prevalence of physical inactivity activity among adults aged 18+ years, age-standardized estimate %	*WHO.* The Global Health Observatory. Available at: *https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-insufficient-physical-activity-among-adults-aged-18-years-(age-standardized-estimate)-(-)*	15/05/2025	0%
Tobacco consumption	Share of population who are daily smokers in individuals aged 15 years old or over, age-standardized percentage of population	OECD Data Explorer. Available at: *https://data-explorer.oecd.org/vis?lc=en&tm=tobacco&pg=0&snb=217&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_HEALTH_LVNG%40DF_HEALTH_LVNG_TC&df[ag]=OECD.ELS.HD&df[vs]=1.0*	15/05/2025	8.13%

7.3. Supplementary Figure 1. Importance of variables in defining the clusters measured by the Mean Decrease Gini (a) and the Mean Decrease accuracy (b)

a)



b)



7.4. Supplementary Table 2. Top five diseases driving cluster separation in the random-forest analysis

Cluster	Disease	Mean Decrease Accuracy	Mean Decrease Gini
1	Age related and other hearing loss	11.20	2.67
1	Lower respiratory infections	8.41	1.59
1	Hypertensive heart disease	7.80	1.33
1	Ischemic heart disease	7.18	1.27
1	Alzheimer’s disease and other dementias	7.10	1.24
2	Other musculoskeletal disorders	9.86	2.02
2	Ischemic stroke	7.07	1.36
2	Age related and other hearing loss	6.85	1.33
2	Alcoholic cardiomyopathy	7.28	0.93
2	Hypertensive heart disease	6.28	0.87
3	Neonatal preterm birth	8.70	1.41
3	Congenital heart anomalies	8.39	1.23
3	Tracheal bronchus and lung cancer	7.73	1.13
3	Diabetes mellitus type 2	4.66	0.73
3	Lower respiratory infections	4.67	0.59

7.5. Supplementary Table 3. Age-standardised DALY rates for the 24 leading diseases across the three epidemiological clusters identified among OECD countries

Disease	Mean (SD)			Median (IQR)			
	Cluster 1 (N = 25)	**Cluster 2** (N = 8)	**Cluster 3** (N = 5)	**Cluster 1** (N = 25)	**Cluster 2** (N = 8)	**Cluster 3** (N = 5)	**p-value**[3]
Age-related and other hearing loss	313.01 (46.08)	451.54 (18.36)	429.32 (37.24)	296.92 (284.05, 344.23)	452.31 (446.44, 458.48)	425.66 (421.92, 444.13)	<0.001
Alcohol use disorders	287.77 (102.93)	489.28 (128.53)	201.67 (89.51)	288.22 (224.83, 350.31)	528.06 (355.35, 567.45)	219.77 (219.49, 225.60)	<0.001
Alcoholic cardiomyopathy	17.83 (16.14)	105. 67 (171.42)	5.54 (4.58)	14.45 (6.89, 20.49)	107.58 (41.57, 189.56)	5.50 (2.34, 7.39)	<0.001
Alzheimer's disease and other dementias	436.16 (46.43)	390.24 (6.67)	373.03 (72.78)	431.61 (418.72, 463.63)	390.23 (384.30, 394.14)	358.18 (357.59, 358.35)	0.001
Anxiety disorders	649.89 (151.08)	411.37 (39.18)	527.72 (83.28)	653.55 (578.47, 756.65)	398.14 (387.71, 422.15)	522.59 (492.89, 553.93)	<0.001
Asthma	259.76 (84.08)	171.19 (84.44)	209.68 (65.65)	256.57 (187.76, 303.77)	135.18 (126.32, 179.75)	249.13 (155.17, 260.15)	0.011
Chronic obstructive pulmonary disease	419.32 (143.49)	339.81 (187.32)	524.67 (215.04)	414.05 (308.70, 498.95)	280.86 (232.71, 375.62)	523.26 (389.90, 538.04)	0.064
Colon and rectum cancer	334.72 (56.62)	463.15 (102.47)	288.95 (56.40)	330.30 (296.00, 365.39)	424.22 (380.69, 549.17)	302.07 (249.10, 312.44)	<0.001
Congenital heart anomalies	94.34 (16.85)	109.10 (41.27)	347.89 (129.66)	94.71 (87.30, 100.71)	109.47 (68.81, 145.53)	398.76 (288.90, 402.60)	0.002

Diabetes mellitus type 2	484.43 (148.57)	615.36 (113.63)	1,154.56 (656.86)	443.16 (388.53, 507.95)	594.05 (517.54, 715.27)	1,012.60 (744.84, 1,017.70)	<0.001
Hypertensive heart disease	90.67 (66.13)	379.11 (302.27)	208.26 (107.44)	66.60 (39.87, 116.49)	255.94 (212.78, 426.98)	162.07 (147.23, 203.60)	<0.001
Ischemic heart disease	906.89 (285.15)	2,408.62 (872.88)	1,607.97 (663.41)	802.23 (721.00, 1,096.13)	2,566.33 (1,833.43, 3,050.33)	1,449.87 (1,166.89, 2,096.66)	<0.001
Ischemic stroke	312.34 (66.03)	838.56 (343.11)	420.79 (218.73)	291.75 (269.59, 344.22)	811.09 (562.40, 1,005.56)	327.15 (289.32, 420.46)	<0.001
Low back pain	1,086.39 (133.89)	1,353.41 (152.85)	931.05 (111.17)	1,079.93 (1,020.10, 1,150.36)	1,367.70 (1,200.00, 1,463.18)	921.30 (849.20, 996.67)	<0.001
Lower respiratory infections	216.42 (82.50)	389.73 (136.05)	556.12 (166.37)	193.50 (158.55, 276.58)	348.63 (290.49, 508.77)	527.56 (452.29, 564.04)	<0.001
Major depressive disorder	574.71 (139.43)	431.81 (127.59)	509.02 (132.36)	552.66 (508.03, 669.94)	395.54 (350.69, 551.87)	552.15 (491.17, 602.84)	0.077
Migraine	627.30 (87.29)	518.86 (18.51)	528.91 (52.43)	649.96 (609.19, 659.74)	519.35 (516.11, 527.51)	532.66 (529.34, 541.54)	0.002
Neck pain	273.19 (68.86)	237.31 (12.87)	271.82 (60.04)	299.14 (237.18, 303.15)	233.09 (232.33, 233.39)	248.60 (247.92, 285.18)	0.12
Neonatal preterm birth	270.82 (91.46)	269.76 (112.67)	659.31 (106.08)	248.52 (201.92, 314.53)	219.89 (183.06, 382.77)	663.65 (606.54, 696.81)	0.002
Opioid use disorders	262.50 (293.21)	179.37 (176.16)	54.48 (21.99)	177.56 (100.51, 333.94)	111.39 (59.52, 243.29)	44.59 (44.39, 55.20)	0.008

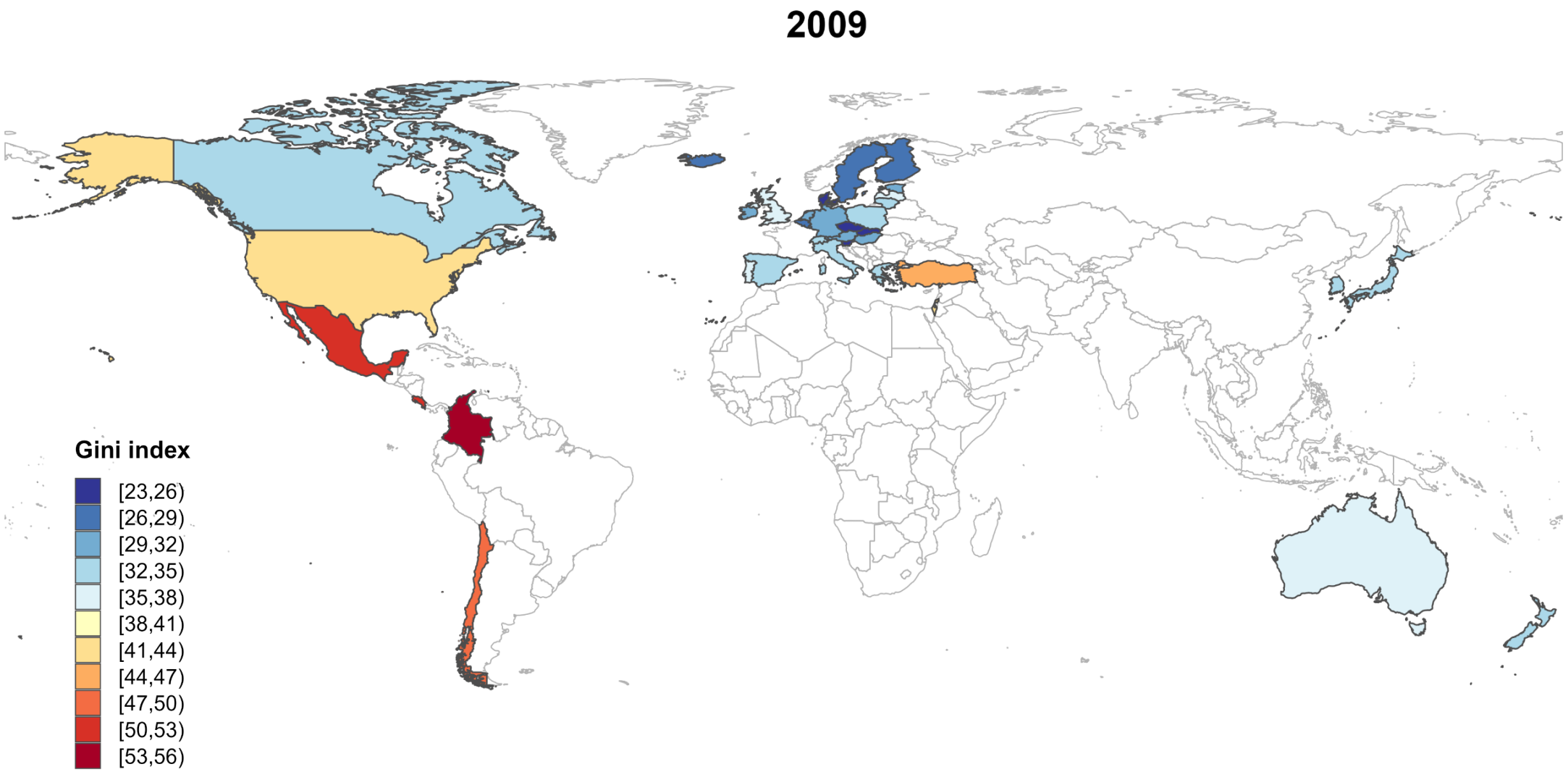
Other cardiovascular and circulatory diseases	128.44 (73.73)	124.61 (58.84)	146.76 (26.07)	101.14 (80.13, 166.13)	107.24 (80.25, 168.88)	148.68 (127.64, 164.83)	0.5
Other gynecological diseases	279.56 (99.54)	171.68 (83.95)	215.10 (113.84)	330.85 (228.80, 345.96)	138.25 (108.64, 259.12)	155.66 (143.73, 222.87)	0.031
Other musculoskeletal disorders	468.16 (191.88)	197.98 (24.49)	745.78 (153.37)	398.86 (365.45, 421.25)	204.95 (172.89, 214.14)	777.95 (777.40, 793.08)	<0.001
Tracheal, bronchus, and lung cancer	597.89 (121.45)	739.27 (216.77)	359.60 (289.48)	596.68 (500.72, 663.12)	654.77 (586.58, 849.38)	234.65 (179.41, 349.81)	0.028

[1]Mean (SD)
[2]Median (Q1, Q3)
[3]Kruskal-Wallis rank sum test

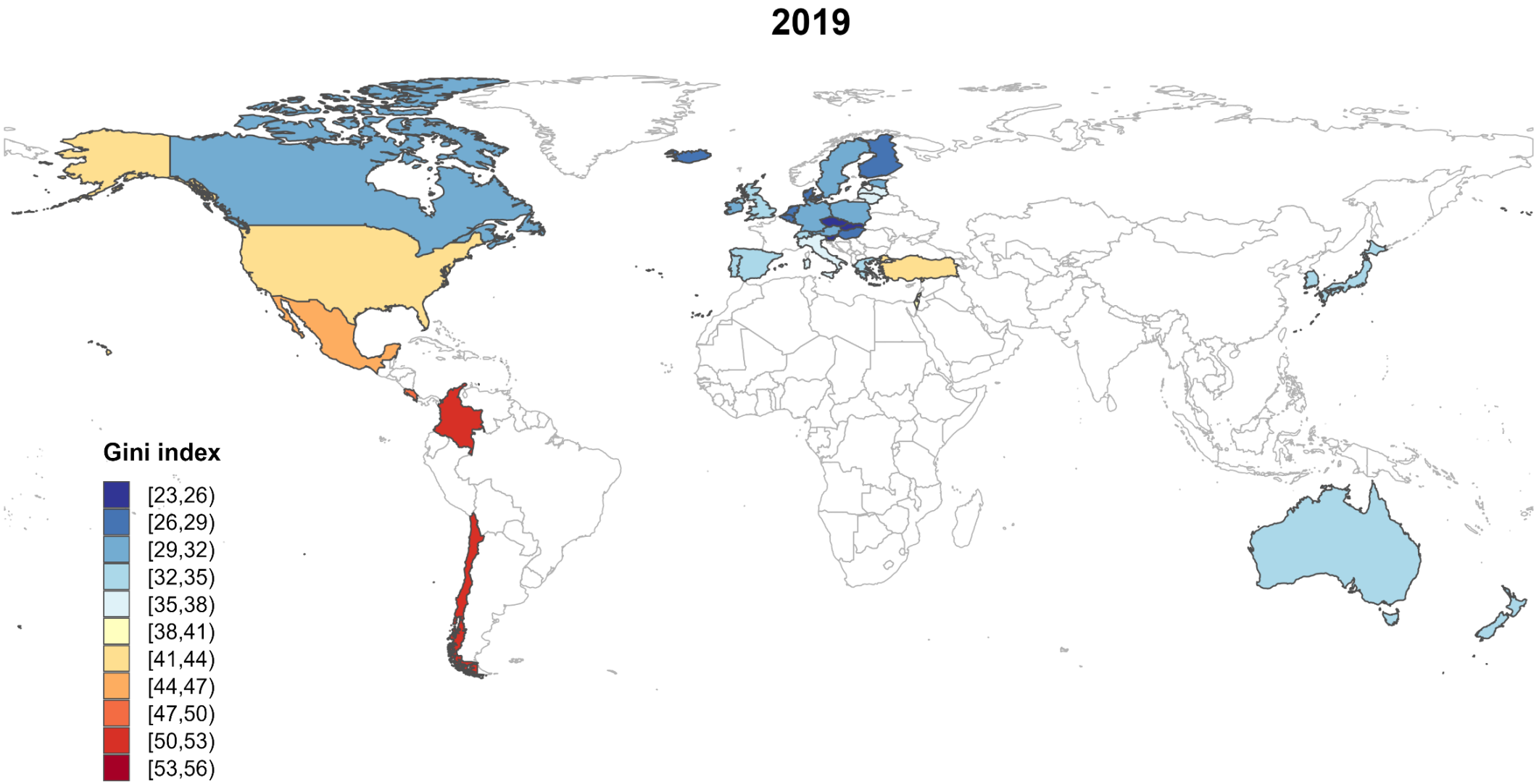
7.6. Supplementary Table 4. List of countries according to their cluster

Cluster	n	Countries
1	25	Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States
2	8	Czechia, Estonia, Hungary, Latvia, Lithuania, Poland, Slovak Republic, and Slovenia
3	5	Chile, Colombia, Costa Rica, Mexico, and Türkiye

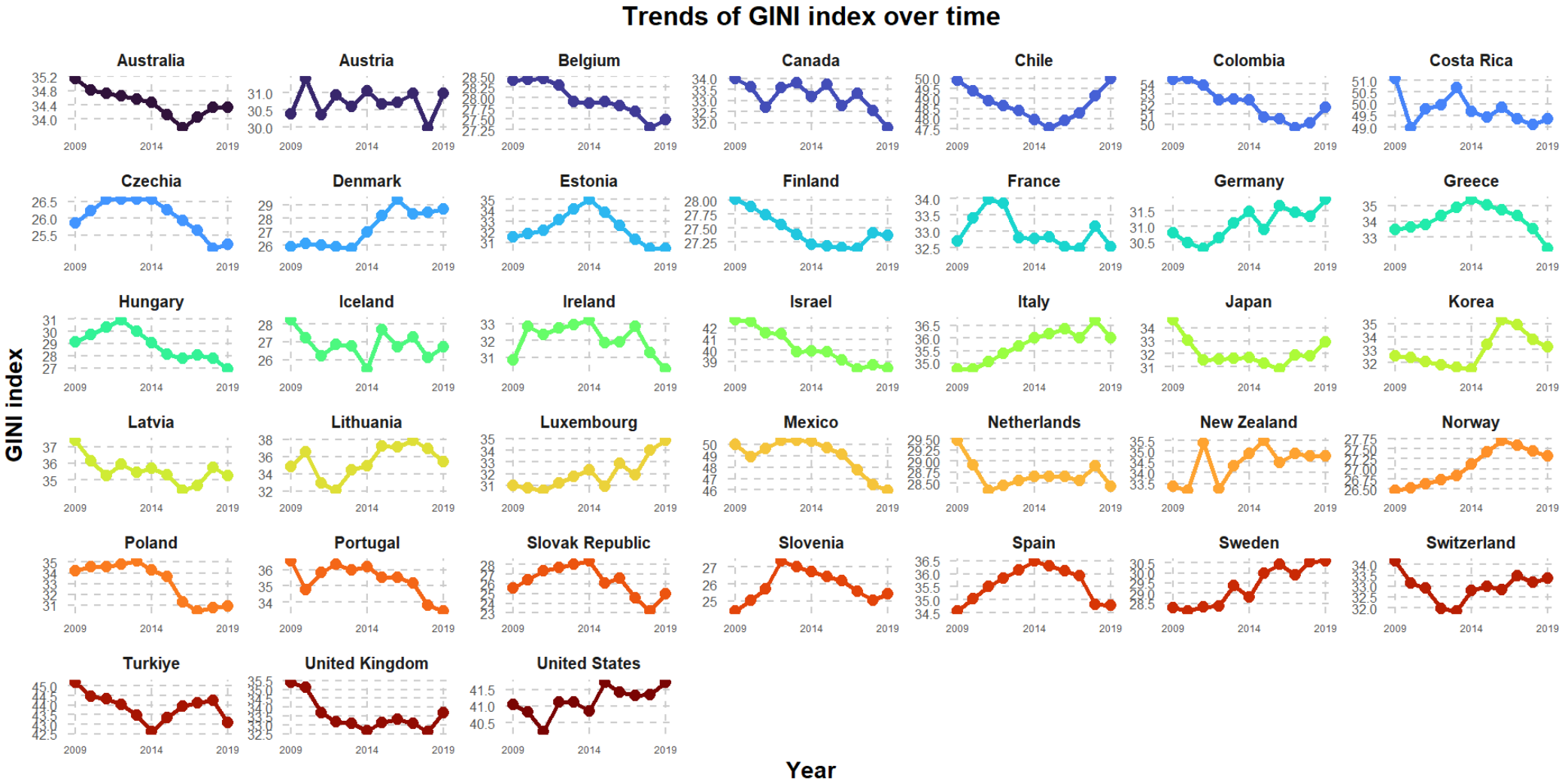
7.7. Supplementary Figure 2. World Map of Gini index in 2009 with OECD countries coloured



7.8. Supplementary Figure 3. World Map of Gini index in 2019 with OECD countries coloured



7.9. Supplementary figure 4. Evolution of the Gini index from 2009 to 2019



7.10. Supplementary Table 5. Descriptive information on socioeconomic indicators and care services in 2009 and 2019 in the 38 OECD countries

Country	**Healthcare expenditure** (% of the GDP per capita)		**Gini index**[a]		**Hospital beds** (per 1,000)		**Physicians** (per 1,000)		**Public Health coverage** (% of expenditure on health)		**Unemployment** (% of unemployed labour force from 25 to 64 years)	
	2009	2019	2009	2019	2009	2019	2009	2019	2009	2019	2009	2019
Australia	8.54	10.21	35.15	34.34	3.77	3.84	3.12	3.83	69.30	71.94	4.43	3.74
Austria	10.23	10.49	30.39	30.98	7.68	7.19	4.70	5.29	75.10	75.07	4.51	4.03
Belgium	10.32	10.76	28.41	27.49	6.15	5.58	2.92	3.16	76.21	75.18	6.59	4.65
Canada	10.70	11.06	34.01	31.76	2.80	2.52	2.04	2.44	70.27	69.50	7.09	4.84
Chile	7.86	9.31	49.92	50.01	2.27	2.03	1.03	2.65	59.03	61.02	7.81	6.41
Colombia	7.31	7.78	54.42	51.67	1.54	1.77	1.62	2.30	74.38	76.45	9.36	9.04
Costa Rica	7.86	7.22	51.09	49.35	1.24	1.10	2.20	2.34	71.73	73.94	5.22	8.49
Czechia	7.25	7.61	25.86	25.22	7.11	6.63	3.58	4.12	83.27	84.87	5.87	1.86
Denmark	10.67	10.15	25.92	28.67	3.50	2.59	3.65	4.27	84.46	83.71	5.20	4.15
Estonia	6.85	6.77	31.52	30.46	5.37	4.53	3.28	3.47	77.01	74.45	12.21	3.99
Finland	9.16	9.17	28.02	27.37	6.25	3.35	3.13	3.57	77.52	77.87	6.52	5.38
France	11.29	11.10	32.70	32.53	6.67	5.82	3.39	3.31	76.51	83.30	7.50	7.34
Germany	11.24	11.72	30.79	31.88	8.24	7.91	3.65	4.39	83.33	84.00	7.41	2.91
Greece	9.41	8.20	33.48	32.27	4.93	4.15	5.82	6.25	68.32	61.51	8.48	16.53
Hungary	7.22	6.28	29.10	26.87	7.14	6.91	3.03	3.49	68.30	68.67	8.81	2.85
Iceland	8.78	8.64	28.24	26.71	3.69	2.80	3.65	3.89	81.64	82.93	5.65	3.03
Ireland	10.52	6.72	30.85	30.29	2.83	2.89	4.04	3.34	77.07	74.25	10.10	4.06

Israel	6.90	7.23	42.71	38.38	3.19	2.97	3.53	3.54	61.22	64.82	6.75	3.36
Italy	8.95	8.66	34.76	36.02	3.66	3.16	3.76	4.05	78.31	73.75	6.53	8.90
Japan	8.96	10.97	34.61	32.94	13.62	12.80	2.21	2.60	81.29	83.96	4.85	2.38
Korea	5.90	8.13	32.53	33.22	8.19	12.43	1.95	2.46	56.89	59.35	3.42	3.37
Latvia	6.10	6.64	37.38	35.27	6.74	5.42	3.18	3.26	59.65	60.12	15.82	6.08
Lithuania	7.37	6.98	34.85	35.42	7.18	6.35	3.83	4.48	72.46	66.38	12.23	6.03
Luxembourg	6.95	5.47	30.98	34.83	5.46	4.26	2.71	2.99	85.75	84.96	4.00	4.50
Mexico	5.85	5.30	50.00	45.99	1.02	0.95	2.00	2.47	46.67	49.17	4.32	2.86
Netherlands	9.99	10.14	29.47	28.42	4.21	3.02	2.92	3.75	83.18	82.80	2.77	3.47
New Zealand	9.62	9.07	33.37	34.78	2.41	2.54	2.58	3.39	81.06	79.65	4.12	2.87
Norway	9.03	10.44	26.48	27.33	4.52	3.48	4.05	4.97	84.44	85.72	2.05	2.85
Poland	6.59	6.46	34.23	30.92	6.65	6.11	2.15	3.26	71.65	71.78	6.84	2.77
Portugal	10.13	9.51	36.53	33.46	3.37	3.49	3.85	5.45	63.86	60.85	9.05	5.82
Slovak Republic	7.95	6.92	25.57	24.99	6.51	5.76	3.30	3.57	73.49	79.79	10.53	5.15
Slovenia	8.54	8.49	24.41	25.43	4.60	4.43	2.40	3.22	73.07	72.79	5.17	4.24
Spain	9.11	9.14	34.58	34.80	3.16	2.95	3.60	4.40	75.07	70.51	15.83	12.86
Sweden	8.77	10.83	28.28	30.57	2.76	2.07	3.75	4.29	82.51	85.12	5.98	5.14
Switzerland	10.14	11.43	34.14	33.37	5.10	4.59	3.84	4.35	64.08	67.81	3.50	4.02
Turkiye	5.49	4.36	45.20	43.09	2.62	2.88	1.64	1.93	80.50	77.69	12.02	11.83
United Kingdom	9.85	9.95	35.41	33.69	3.27	2.45	2.65	2.94	82.35	79.24	5.72	2.75
United States	16.20	16.55	41.04	41.73	3.08	2.80	2.57	3.52	48.34	82.73	8.14	3.17

[a] Gini index is a measure of inequality based on household income. It ranges between 0 and 100, where 0 indicates total equality and 100 indicates total inequality

7.11. **Supplementary Table 6.** Spearman correlation between the studied covariates

	Gini Index	***Healthcare expenditure***	**Hospital beds**	**Physicians**	**Public healthcare coverage**	**Unemployment rate**	**Obesity**	**Tobacco**	**Physical inactivity**
Gini Index	1								
Healthcare expenditure	**-0.26*** **	1							
Hospital beds	**-0.47*** **	0.00	1						
Physicians	**-0.36*** **	**0.32*** **	**0.23*** **	1					
Public healthcare coverage	**-0.51*** **	**0.33*** **	0.06	**0.16** **	1				
Unemployment rate	**0.25*** **	**-0.13***	-0.06	0.04	**-0.26*** **	1			
Obesity	**0.44*** **	**-0.32*** **	**-0.43*** **	**-0.29*** **	**-0.31*** **	**0.103***	1		
Tobacco	-0.02	**-0.15** **	**0.49*** **	0.08	**-0.15** **	**0.43*** **	**-0.12***	1	
Physical inactivity	**0.43*** **	-0.03	**-0.15** **	**-0.27*** **	**-0.26*** **	**0.11***	**0.13** **	0	1
Hypertension	**-0.19*** **	**-0.43*** **	**0.26*** **	-0.02	-0.02	**0.34*** **	**0.12***	**0.34*** **	**-0.23*** **

Rho from Spearman correlation are shown. Bold values express significant results. * p<0.05, ** p<0.01, *** p<0.001

7.12. Supplementary Table 7. Country and year fixed effect crude model

	Model 1	Model 2	Model 3
Variable	**IHD**	**Neonatal preterm birth**	**Anxiety disorders**
Gini index	-0.012 (0.006) *	0.018 (0.003) ***	-0.0003 (0.001)
R-squared	0.057	0.0761	0.136
Hausman test P-value	<0.001	<0.001	<0.001

Note: The cluster robust standard errors are used in the estimation[7]. SEs in parentheses; *** P ≤ 0.001, ** p ≤ 0.01, * p ≤ 0.05.

[7] Since the observations of the same country in different years presented serial autocorrelation, and the assumption of standard error requires that the standard error is independent and identically distributed, the estimation of common standard error is not accurate, so this study reports the cluster robust standard error.

7.13. Supplementary Table 8. Country and year fixed effect model using with Palma ratio

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	IHD	IHD and risk factors	Neonatal preterm birth	Neonatal preterm birth and risk factors	Anxiety disorders	Anxiety disorders
Palma ratio	-0.175 (0.068) **	-0.141 (0.051) **	0.158 (0.022) ***	0.116 (0.041) **	-0.026 (0.026)	-0.017 (0.011)
Healthcare expenditure	-0.028 (0.013) *	-0.022 (0.010) *	-0.013 (0.006) *	-0.016 (0.015)	0.004 (0.005)	0.006 (0.001)***
Beds	0.029 (0.017)	0.017 (0.021)	-0.019 (0.009) *	0.002 (0.019)	-0.05 (0.03)	0.000 (0.003)
Physicians	0.073 (0.028*)	0.034 (0.028)	0.017 (0.010)	0.037 (0.036)	-0.018 (0.021)	-0.013 (0.006)*
Public health coverage	0.004 (0.01) ***	0.003 (0.001) ***	0.002 (0.001) **	0.003 (0.001)	-0.001 (0.0001)*	-0.001 (0.001)
Unemployment rate	0.002 (0.03) **	0.001 (0.02)	0.007 (0.002) **	0.007 (0.004) *	0.002 (0.001)*	0.002 (0.001)**
Hypertension		-0.020 (0.010)		-0.007 (0.015)		0.005 (0.002)*
Physical inactivity		0.006 (0.04)		-0.006 (0.006)		-0.003 (0.001)***
Tobacco		0.010 (0.003) ***		-0.007 (0.004) *		0.001 (0.001)
Obesity		0.020 (0.012)		-0.003 (0.011)		0.003 (0.002)*
R-squared	0.253	0.388	0.136	0.193	0.139	0.189
Hausman test P-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Note: The cluster robust standard errors are used in the estimation . SE in parentheses; *** p ≤ 0.001, ** p ≤ 0.01, * p ≤ 0.05.